TIME-VARYING MARKOV MODELS

OF SCHOOL ENROLMENT

by

MARIA MANUELA MARTINS PINHEIRO DE MAGALHAES

(Volume I)

A Thesis submitted to the Department of Economics and Management Science at the University of Keele for the degree of Doctor of Philosophy

University of Keele,

July 1987

The following have been excluded at the request of the university:

The quotes on pgs 14, 15, 16, 17, 23 & 32.

ABSTRACT

This study uses Markov models to develop a general quantitative approach to aid the modelling of school enrolment. The performance of the stationary Markov chain model, widely used during the 1960s and early 1970s to project school enrolment, has thrown into relief the limitations of the traditional model. Only a few studies have thoroughly tested the model over a period of time to determine whether it is really valid for predictive purposes.

The present study starts by testing the stationary Markov model using data over a twelve year period for a subsystem of the Portuguese educational system, the model being applied to the whole country and to each district into which the country is administratively divided. Several least squares estimation procedures are performed to produce estimates of the transition probabilities. As expected this model proves to be inappropriate, generating biased and non-efficient estimates for the transition probabilities.

Assuming that the non-stationarity of the transition probabilities is due to causal factors, linear behavioural relationships are included in the model. An extended Markov model with time-varying transition probabilities is developed and applied to the same Portuguese educational subsystem. Seventeen explanatory variables, divided into supply-side factors and demand-side factors, are used, and stepwise regression and pooled cross-section time-series regression are performed to produce estimates of the time-varying transition probabilities. Principal components analysis is also applied on supply factors and demand factors and new sets of explanatory variables are used.

The results show that the patterns of the time-varying transition probability estimates describe reasonably well the patterns of the corresponding observed point estimates. This suggests that it is appropriate to include a causal structure in the model. Having established the causal relationship influencing the time-varying transition probabilities, an analysis of these relationships suggests both policy implications of this work and areas for future research.

ACKNOWLEDGMENTS

I wish to thank all those who have helped me during the three years that I have worked in this research project. I wish to express my sincere thanks to my supervisor Dr. J. Proops. I am deeply indebted to him for his invaluable advice, constructive criticism, generous help and guidance. Moreover my thanks are also extended to the encouragement and constant support of Professor L. Fishman, Head of the Department of Economics and Management Science and to Dr. L. Rosenthal for his stimulating suggestions. Further, my sincere thanks and appreciation to the department secretaries, especially Mrs. S. Beech and the staff of the Computer Centre, particularly Dr. P. Collis for his invaluable advice and guidance in the computer programming. Also the staff of the University Library for their assistance.

I gratefully acknowledge the Portuguese governmental department JNICT for granting me a NATO fellowship, essential financial support for the development of this thesis. I am also thankful to INIC for providing me with supplementary financial support during my first year of stay at Keele University. Acknowledgment is also extended to ISCTE and to the Planning Bureau (GEP) of the Ministry of Education for giving me research leave that allowed me to stay at Keele and made the completion of this thesis possible.

I am indebted to Mrs. L. Smith who typed most of the tables.

LIST OF CONTENTS

Chapter 1

INTRODUCTION

> "Educational planning is particular prone to uncertainty about the future, since even the present relationship between the supply of qualified students and the demand for educated people from industry and government is little understood. It is advisable to build into the system, the kind of flexibility that allows it to adjust automatically to bottlenecks and surpluses. Educational planning should be characterized by a multiplicity of alternatives in producing and utilizing educated manpower." [BLAUG, 1967; p.273]

This dissertation develops a general quantitative approach to the better modelling and understanding of the determinants of promotion, repetition and drop-out for school enrolment. The feasibility of this approach is explored using data covering a twelve year period (1970/71 - 1981/82) for seven grades of the state schools (basic preparatory and secondary levels of education) of the Portuguese educational system. In Portugal, as in other European countries, at the end of each academic year a student needs to reach a certain level in order to progress to the next grade, otherwise he/she fails and either repeats the grade, or leaves the system and is considered as a drop-out. Thus the promotion of a student to the next grade is not made automatically.

Various approaches to the study of school enrolment, such as Markov methods, regression analysis and simulation, have been used since the early 1960s, to produce projections of the number of students enrolled, graduates and drop-outs. These approaches had the aim of providing the educational planners and decision-makers with information for the preparation of educational plans and policy implementation. However, to be certain that a model is really valid for predictive purposes, it must describe with a reasonable confidence the historical trend of observed values. Scarcely any studies, however, have tested their models to determine if they generate the historical trend.

Among the approaches presented in the literature, the Markov chain approach appears to be the most widely applied, as it is appropriate in describing the movements of students within an educational system such as that of Portugal. However, all the stochastic approaches used to study the behaviour of school enrolment using Markov processes, assume that the educational system can be described by a stationary Markov chain, that is, the transition probabilities are constant over the sample periods. Stochastic models applied to school enrolment have, therefore, been entirely passive in terms of behaviour. This is a strong assumption when dealing with human beings, and has been widely criticized in the literature, as students' behaviour may vary according to causal factors, such as family characteristics, community characteristics, school and economic characteristics, for instance.

Qualitative analysis became, therefore, an important element in planning for future developments of the education system. Several studies have been carried out in order to analyse the factors affecting the demand for education. However, not much has been done in the area of simultaneous estimation of time-varying transition probabilities. Extending the traditional stationary Markov chain model of school enrolment by allowing the transition probabilities to be affected by changes in causal factors is the main purpose of this thesis.

The first part of the study tries to demonstrate the weakness of the assumption of time invariant transition probabilities; it starts by describing the "basic Markov model" and then applies it to the subsystem of the Portuguese educational system selected in this study.

In a second part, the restriction of time invariant transition probabilities is dropped in favour of a more flexible assumption, allowing the transition probabilities to be function of a set of explanatory variables. A new "extended Markov model" is developed and applied to the same subsystem of the Portuguese educational system.

The study starts by presenting in Chapter 2 an overview of the different approaches to educational planning and a survey of the relevant literature. Chapters 3-5 are concerned with the basic Markov model. First, the theoretical framework of the model and the different methods of estimating the transition probabilities are described in Chapter 3.

This is followed in Chapter 4 by the application of the basic Markov model to the basic preparatory and secondary levels of the Portuguese educational system. The analysis undertaken in this chapter applies to the whole country and the least squares estimation procedure (unrestricted and restricted) is used to produce estimators for the different transition probabilities. Assuming that certain disturbances in the data are due to the return of students from the old colonies Angola and Mozambique after the revolution that took place in April 1974, an iterative process is applied in order to separate these students from the observed data and to obtain a "corrected data matrix". The least squares estimation procedure is applied to calculate new estimators for the transition probabilities.

A regional level application of the basic Markov model is performed in Chapter 5, with the aim of improving knowledge of the behaviour of the different point estimates of the transition probabilities and their estimators. The same structure of analysis developed in Chapter 4 for the whole country, is applied in Chapter 5 to the eighteen districts into which the country is administratively divided.

Chapters 6-8 are concerned with the investigation of the reliability of the application of the extended Markov model to school enrolment. Chapter 6 is devoted to the formulation of the theoretical elements of the extended Markov model, in which the parameters vary over the sample period. A linear causal structure is incorporated in the model, allowing parameters to become functions of exogenous variables. The chapter describes also the least squares estimation procedures to derive the estimates of the transition probabilities.

Chapter 7 concentrates on the application of the extended Markov model to the basic preparatory and secondary levels of the Portuguese educational system, using the country as a whole. Seventeen explanatory variables, broadly divided into supply factors and demand factors, are included in the model and unrestricted and restricted OLS estimation procedures are performed in order to obtain time-varying estimates for the repetition and promotion probabilities. Due to the existence of multicollinearity between the explanatory variables selected, principal components analysis is performed and regressions on the principal components of the supply side and demand side variables are carried out to produce new estimates for the transition probabilities.

Regional data are used in Chapter 8 in order to improve the results of the application of the extended Markov model to the Portuguese educational system. The data corresponding to the different districts have been stacked and are treated as a single set. Stepwise regressions and pooled cross-section time-series analysis are applied to produce the estimates for the transition probabilities. Principal components analysis is applied to each district individually and new estimation procedures are performed on the new sets of stacked data.

The analysis of the results and policy implications are presented in Chapter 9, where careful attention is paid to the analysis of the behavioural relationship estimated. Finally, conclusion and suggestions for further research are presented in Chapter 10.

The bibliography contains all material referenced in the text and also other material that influenced the writing of this thesis, but which has not been referenced directly.

Chapter 2

THE EDUCATION SYSTEM AND THE ECONOMY: A REVIEW OF THE METHODOLOGIES IN THE LITERATURE

2.1. An Overview of the Educational Planning Process and the Factors Affecting its Evolution

The 'human investment revolution in economic thought' initiated by SCHULTZ [1961] led some economists to take an interest in education. From then the economics of education became an economic field in its own right, and its importance in academic research, as well as in socio-economic planning, has continued to grow. Education has come to be viewed as one of the main factors influencing economic growth. The approach to education was changed from viewing it as a form of consumption, or as a device for the transmission of culture; instead, spending on education came to be seen as an investment in man, which results in an improvement in the quality of the labour force and corresponding increases in the productivity of labour. Increased productivity results in an increase in personal earnings and in the economic growth of the country. Work by DENISON [1962], who measured the contribution of formal education to economic growth in the context of an aggregate production function, and by SHULTZ [1961], who studied the United States growth in the 20th century, revealed a growing interest in this view. These authors claimed that twenty per cent of the growth rate in the study period was due to an increase in the education level of the labour force. The governments of the countries of the world were then convinced of the economic benefit of education and the concept of 'human capital' has spread into countries' development plans.

Since education was considered to be an important factor of economic growth, a large amount of society's resources was devoted to education each year, necessitating a method of allocating these resources efficiently to the educational sector. Planning has become more important due to its information role, upon which decision makers will rely to allocate financial and other resources. The growth of interest in educational planning was then remarkable. Most of the Ministries of Education around the world were engaged in short-term and long-term education plans, drawn up as a part of social and economic development.

The basic goals of education policy throughout the 1960s were then:

- meeting the manpower needs of the economy;
- ensuring equal educational opportunities to all citizens;
- enhancing efficiency.

The first goal received high priority because the national authorities had felt that without an adequate number of well qualified workers, the objective of economic growth would have been seriously compromised. The second goal derives from the political agreement to offer equal opportunities to the citizens, and the assumption that this could be met through the generalised democratisation of educational opportunities. The third goal offered a guarantee that the achievement of the other two goals would be pursued under the condition of economic efficiency.

Three main approaches have been applied to educational system planning to provide decision makers with information on how well the education system is achieving its objectives. All these approaches may give different answers and are generally considered competitive; namely the 'manpower requirements', the 'social demand', and the 'rate-of-return' approaches. In context, however, they can be complementary to each other.

The manpower requirements approach was introduced by PARNES [1962] and was widely used in many countries. The method projects the occupational and educational composition of the labour force to some future date on the basis of estimated sectoral growth rates and productivity changes, and then translates these projections into required educational system supply output. The aim of this method of approach was to reduce the unemployment and underemployment due to the incompatibility of the skill structure and the manpower requirements. Also, by avoiding structural imbalances, it can stimulate economic growth. Planning using the manpower requirements approach is thus based on the hypothesis that the skill structure of the labour force is strictly determined by production.

The social demand approach projects the growth of the educational system by extrapolating recent trends into the future, according to predetermined educational targets. Planners determine desired student enrolment growth rates, student places, the number of teachers and other factors, in order to meet society's demand for education. It must be noted that no optimisation is applied. The existing educational systems, presumed to be in equilibrium, are simply expanded in the traditional manner until the target is reached. This approach aims only to satisfy private demand for education as a main social and political objective of the country.

The 'rate-of-return' approach has the goal of investing resources in those types and levels of education yielding the highest returns. Planning decisions are then based either on the computed internal rate-of-return or on the net present value of those benefits and costs associated with the schooling. This approach, designed to handle the problem of efficient resource allocation, is based on economic analysis more than any of the other approaches already described. According to BLAUG [1972], the advantage of this approach is that it leads to optimising cost-benefit comparison patterns, paying attention, therefore, to the internal efficiency of both the educational system and the labour market. The rate-of-return approach takes into account the employment benefits stemming from every additional year of study over the employment life span of an individual.

The rate-of-return approach was rarely used for public educational decisions [see SOUMELLIS, 1981]. It was the other two approaches which mainly dominated educational planning, particularly during the 1960s.

The implementation of these approaches led to two extreme notions of educational planning: 'technocratic' educational planning and the 'political' mode of planning. The technocratic model implies a clear distinction between the policy-maker and the planner. The policy-maker (or policy-making group) establishes a series of strategic policy objectives (social equality, meeting national needs for qualified manpower) so that the planner (or planning team) expresses these objectives in operational terms. This dual nature of planning was used in many countries (including Portugal) and is still advocated by several specialists. The political model, on the other hand, denies the existence of policy-makers who establish the strategic objectives. Instead, policies are made on the basis of a series of tactical decisions which result from pressure groups, of which no single group is powerful enough to be able to impose its views on the others. The planner is in this case a mediator between the different interests involved. It is the planner who has to be well informed in order to establish a policy which will be the result of the various tactical pressures. This kind of model denies the prospective function of educational planning and focuses only on short-term policies, particularly linked to the annual allocation of resources through the budget.

In technocratic educational planning, given the two major policy objectives, i.e. social equality and economic growth, several specific techniques have been developed, with the aim of accommodating the social demand and satisfying the manpower requirements. However, in its early years, the statistical system was very poor, the statistics of education being inadequate for the analysis of educational development. It is important to note that VAIZEY and EDDING [in WILLIAMS (1979), p.6] undertook what were essentially data gathering exercises, especially for areas that investigate interdependencies of economic and educational developments in order to obtain long-term forecasts. The OECD and UNESCO, as international organisations, also played an important role in the development of educational statistics [see SVENNILSON et al. (1962), OECD (1967b) and UNESCO (1955;1966)]. The role of OECD and

UNESCO in supporting or in undertaking development analysis itself, was also very considerable; the effort is being continued [see OECD (1974), UNESCO (1974c; 1976b) and JOHNSTONE (1981)].

The mathematical models developed in the 1960s fall into three main categories:

- models exclusively for the education sector;
- models exclusively for the manpower sector;
- models relating the education sector to other social sectors.

The models exclusively for the education sector have the main objective of calculating the evolution of some variables of the educational system only. Three basic models are in use: (i) student-flow models; (ii) models for calculating the demand for teachers on the basis of the number of students enrolled in each grade; (iii) cost models for calculating the total amount of financial resources needed.

Student-flow models are of two types. One is based explicitly on population figures and observed enrolment ratios. The second uses the population figures at the first grade to calculate the new entrants; after that, the model is based on transition coefficients from each grade of the educational system to the next. Flow models of this second type have been used by Ministries of Education (Portugal is an example) for forecasting the evolution of the enrolment in the school system. These models make up the technical basis of the social demand approach to educational planning.

For the manpower sector exclusively, the aim of the models is to help in forecasting the manpower requirements of the economy, which are going to be used as targets for planning the educational supply by types of graduates at different educational levels. The models employed are of three types: (i) extrapolation of past trends in the growth of an occupation; (ii) models estimating anticipated supply of particular types of worker on the basis of present stocks, anticipated education supply and anticipated losses to death, retirement and withdrawal from the labour market; (iii) regression models correlating the number of employees in a certain occupation with total employment, production, total population, national income, productivity, or other variables.

The models that try to relate the education sector to other social sectors, attempt to link primarily the education sector with the manpower and economic sectors. Two types of models have been developed in this area: (i) input-output models which attempt to link the educational sector (student-flow matrix) to the economic production sector (activity input-output matrix) and the labour market (teachers in the form of an occupational flow matrix); (ii) multi-equation econometric planning models linking the various levels of the educational system to the economic production system.

All OECD countries which engaged, in the early 1960s, in some kind of educational planning followed the technocratic model.

Meeting the equality objective was interpreted, in the 1960s, as encouraging social demand for all types of education and satisfying it through a continuously growing educational system. Meeting the economic growth objective implied estimating the manpower requirements for attaining desired growth targets on the basis of two assumptions: a given macro-production function and a desired productivity growth. The mathematical models were the basis for the planning operations to achieve these two main educational policy objectives. However, these two main policy objectives are not necessarily compatible. Their priorities differ from country to country, depending on the country's stage of development.

Within the OECD contribution, the relatively less economically developed countries put their efforts into manpower planning, whereas the group of more developed countries followed a social demand approach. The first group participated in the Mediterranean Regional Project (MRP) (Greece, Italy, Portugal, Spain, Turkey and Yugoslavia) whilst the countries of the second group cooperated within the Educational Investment Planning (EIP) (Ireland, Sweden, Netherlands, Austria).

In the early 1970s, the most dramatic development in the economics of education was the increasingly critical reevaluation of the promises of the educational planning established after the introduction of the 'human capital' concept. The enrolment explosion that took place all over the world since 1945 began to slow down, destroying the optimism that the expansion of education would effectively equalise life chances in industrialised societies. This optimism of economists was then severely shaken in the late 1970s. Economists became aware of the difficulties of finding the link between education and income, and came to recognise, therefore, the limits of education as an instrument for equalising income and promoting economic growth.

Thus, despite its appeal, the human capital concept has been shown to have many defects that limit its application to the measurement of the social value of education. In fact, even if education and earnings are positively related for individuals, this does not mean that is so for the society as a whole. Education might not contribute anything to growth in total income but might act as a powerful agent of income redistribution, favouring people with higher levels of education [ARROW, 1973]. Even if one overcomes the obstacles of accurate calculations of human capital, this concept would be an imperfect guide to policy. One is still left with the problem of whose satisfaction should be maximised. Only value judgements could answer this question. As MARSHALL, BRIGGS and KING [1984] pointed out:

But BLAUG [1976] was probably right when he asserted:

The rapid accumulation of models during the 1960s was followed by the criticism of certain methods. There was a change in the perception and practice of educational planning 'from the quantitative technology of planning towards interactive, consultative and participative deliberation; from convergent planning in which authoritative allocation can be decided towards multi-value assessments which will hold open a larger number of options' [KOGAN (1980), p.1]. But this does not mean that quantitative analysis is unnecessary; on the contrary, detailed quantitative planning is essential. However, several recommendations were made that educational planning be based on politico-sociological studies on how, and how far, traditional educational patterns have contributed to the failure of social and economic progress in the past. Micro analysis became then of great relevance, more sophisticated approaches to examine the educational issues being suggested, proposing interdisciplinary work with sociologists and psychologists of education. Even during the so called 'golden period' several remarks on the different educational planning approaches have emerged; BLAUG's quotation displayed in the introduction of the present study is an example of such remarks.

In a review of STONE's work, BRAY [1965] made the important point that there are dangers in providing background information which was then used somewhat mysteriously, to reach actual decisions:

A survey of the existing mathematical models for the education sector was carried out by the OECD [see OECD, 1973]. Practically all the member countries participated (20 countries) and a detailed questionnaire was used. The survey confirms the manpower requirements approach to be the governments' preference for educational planning, also showing that 36% of the models covered higher education only. For the education system represented by itself, only <u>limited attention was paid to examining the behaviour</u>

which governs the values of the transition coefficients. The results of the survey were presented for review to a meeting of experts in the field, who were asked to make recommendations as to the best way of pursuing work in educational model-building. One of the suggestions presented was the construction of more complex models which specify simultaneously interdependent equation systems, in order to reflect the complexities of the real situation in education. Therefore, student parameter models which attempt to evaluate and explain the parameters of student forecasting models and models concerned with productivity or student attainment, which evaluate the influences of various inputs into the educational system on its outputs, should be further explored. A straightforward abstract of their recommendations follows:

As the remarks and comments show, up to the end of the 1960s and early 1970s, forecasting and planning had not led to a solid basis for plotting long-term trends in the field of education. Educational planning requires an understanding of the function of education as a part of today's social and economic change.

THOOLEN [1970], for example, has shown that, apart from some technical differences, a large degree of similarity exists between the OECD Simulation Option Model (S.O.M.) and a Dutch educational model developed by the Dutch Central Planning Bureau, and partly based on the concept of student flows. Investigating the possibilities and limitations of both models for the purpose of educational forecasting and simulation, he stated:

These recommendations involve significantly deeper social demand analysis and a consideration of alternative educational systems and processes. Suggestions have then been presented to construct more complex models which specify simultaneously independent equation systems, in order to reflect the complexities of the real situation in education. Important auxiliaries to resource requirement models, such as (i) student parameter models, which attempt to evaluate and explain the parameters of student forecasting models, and (ii) models concerned with productivity or student attainment, which evaluate the influences of various inputs into the educational system on its outputs, should be further explored.

The context of future educational policy development of the late 1970s has then been different from the context which existed in the 1960s. The basic characteristics of the late 1970s were:

- 'increasing doubt as to the potential of education per se to help attain the socio-economic objectives and particularly the objective of equality of opportunity given the persistence of social, economic and educational inequalities;

- an economic system which has suffered two consecutive crises, with rapid decrease in economic activity and consequently a serious unemployment problem, with substantial loss of public financing capability;

- because of the economic crisis social priorities have changed, with education losing its previously privileged social rank.' [SOUMELIS (1983), p.27]

A further accentuation of the above trends have been shown during the 1980s so that the majority of OECD countries present, at the moment, the following characteristics:

- 'low economic activity;

- high unemployment rates and particularly youth unemployment;

- high levels of inflation;

- changing attitudes towards the value and roles of education for society, as for the individual resulting in low political priority for education;

- drastically contracting public budgets;

- rapid technological changes in some economic sectors;

- profound demographic changes resulting in decreasing school population and faster-growing population at the retirement age;

- growing demand for different kinds of education from new population groups.' [SOUMELIS (1983), p.27]

In this context, social demand studies, attempting to examine the qualitative aspects of education, and mini-tracer surveys started being performed by the different countries. These studies, while being essential, need to be combined very carefully. Educational planning in the past has been essentially quantitative. A fully effective educational planning process should comprise qualitative as well as quantitative aspects. In recommendations for further studies underlining the educational planning process, WILLIAMS [1983] remarked:

> 'To claim, therefore, that educational planners must be more concerned with the qualitative aspects of education is not to argue for decision-making based simply upon competing value judgements. Methods must be found of bringing matters relating to the context and methods of educational provision into a systematic framework alongside more obviously quantifiable features such as pupil numbers and costs.' (p.354)

2.2. A Review of the Literature of the Different Approaches to Educational Planning

Since the early 1960s, a variety of studies have been undertaken as support to the three main approaches to educational planning: manpower requirements, social demand, and rate of return approaches. There exists a quite wide range of studies about the subject and several manuals and surveys have been published since the early 1960s. All the models represent attempts to explain the nature of the education system as a whole, or of parts of such a system. Some countries have developed systems of analysis employing a number of different models or projection techniques, usually applied by different Ministries or planning departments. The models can be completely separate but their results are needed as inputs to other models, and the different sets of projections are used to identify possible future inconsistencies and imbalances. In other cases, more sophisticated versions are used, employing only a large integrated model, yielding, for example, simultaneous enrolment projections and projections for manpower demand and supply, consistent with stated growth targets for the economy.

2.2.1. The Manpower Requirements Approach

As already mentioned, PARNES's [1962] methodology, described in the OECD's Mediteranean Regional Project, was the one that in the early 1960's became a guideline in the area of manpower planning. The method used starts by an initial projection of a desirable GDP in a future year disaggregated by major sectors. The sectoral GDPs are then disaggregated by industries. An average labour-output coefficient (the reciprocal of the average productivity of the labour) is applied to the sectoral or industrial GDP targets, yielding a forecast of labour requirements by sector of industry. The labour force is then distributed among a number of mutually exclusive occupational categories and the occupational structure of the labour force is converted into an educational structure by applying a standard measure of the level of education required to perform 'adequately' in each occupation.

The difficulties in this method arise essentially in forecasting the labour-output coefficients and in transferring the labour requirements by occupation into labour requirements by educational qualification. Despite many attempts it cannot be said that these difficulties have been solved satisfactorily. The application of this method was, therefore, quite limited: out of the six southern European countries which entered the project, only Spain would use the proposed methodology; for the remaining countries, Portugal being one of them, the project was essentially a starting point for educational planning, the objectives being established according to social demand factors (student places, number of teachers, student enrolment growth rates), rather than based on well defined education needs for economic growth.

The CORREA-TINBERGEN [1962] model is one of the most famous approaches to educational planning. This model is an attempt to relate education and the economy, investigating the balanced growth conditions for smooth educational expansion. One sector of production and three levels of education were represented. The model assumes fixed requirements for each type of manpower per unit of output, as well as fixed teacher-pupil ratios. Each of the three school stages (Primary, Secondary and Tertiary level) were assumed to have a training period of six years, which was used as the time unit of the model. Although this time unit simplifies the analysis, it seems to be too large a time unit to give realistic results. CORREA [1963] has presented a further elaboration of this model, assuming that the development of the output is determined by a growth model of the Harrod-Domar type (fixed capital coefficient and fixed savings ratios). In a later paper TINBERGEN and BOS [1965] suggested some improvements to the model, including non-linear relationships between inputs of educated manpower and Gross National Product, sectoral disaggregation of production and taking into account the drop-outs during the educational process. A shortcoming of this model, however, was that the demand for educated manpower was determined in terms of the required number of workers, without reference to their relative wages, as in the manpower requirements models.

BOWLES [1967] constructed a linear programming model which allows the simultaneous computation of optimal enrolment levels in each type of education, an optimal pattern of import (or export) of educated labour, and the determination of the efficient educational system as producer of labour. BOWLES's study was applied to Northern Nigeria, the objective function being the increment in discounted net life earnings attributable to additional years of education.

A dynamic linear programming model for educational and economic planning was developed by ADELMAN [1966] with Argentinean data. The aim of the study was to determine the optimal extent and composition of resource allocation to education and training. This was attempted by considering educational investment in real capital, which involved the use of both manpower requirements, and cost-benefit educational approaches to educational planning.

In France, research workers at the Research Centre for medium and long-term Economic Forecasting [GIRARDIEU, 1967] have developed a model for the optimal allocation of resources between the productive economy and the educational system. This optimization was reached by the maximisation under constraint of a social preference function. This model represents an attempt to include cultural and educational needs quantitatively within the forecasting of manpower needs.

All these models were developed during the 1960s, which show that, since the discredit that started emerging in the early 1970s, economists have turned their attention to other kinds of analyses, no more significant attention being paid to this approach.

The manpower requirements approach can provide useful insights in cases where the relationships between education and occupation are clear and direct. This is one of the weaknesses of the model as no allowance is made for substitutability between labour of different skills and levels of education. Also, no changes in productivity of labour, or in the technologies which are likely to influence the future, are allowed. The manpower requirements approach could, however, be beneficial if it is used with additional investigations of costs and benefits of education.

The comments on the TINBERGEN - BOS model presented by SEN [1964], reemphasize the weakness of this approach:

2.2.2. The Social Demand Approach

Various rather primitive methods which can be used in school enrolment were surveyed by JACOBI [1959]. One of the methods uses survival ratios, representing survival of students from one school stage to a later one.

ARMITAGE and SMITH [1967] presented a computable model of the British educational system. This computable model was jointly carried out by the Department of Education and Science and the Unit for Economic and Statistical Studies on Higher Education (London School of Economics and Political Science). This model runs the projections of the school population by sex and age in a given year, based upon enrolment in the grade or level below in the previous year, and upon coefficients describing the flows of students between grades and levels from one year to the next (the student flow model, which in its main features is of the Markov chain type).

A projection model particularly adapted to conditions in Asian countries was developed by UNESCO in 1965 and was used to quantify the implications of targets for educational development in three groups of countries in this region. The methodology used in this model has since been revised and published as the Educational Simulation Model (ESM) (this model is computerised and available for use by UNESCO's member States at their request). This is a pure deterministic model, which predicts enrolment in each course of studies, calculates the number of teachers needed by educational level, and determines the recurring and capital costs.

STONE [1965] has also presented a model of an educational system, giving emphasis to the students' demand for education. The system of education was then divided into a number of successive processes: a compulsory process through which all students must pass and, thereafter, voluntary processes, listed in successive age order. The model was treated as a dynamic input-output model, the input being less trained students and the output being students who have gone through an additional process stage. Part of one year's output is next year's input; the rest graduate or leave the system. Noticeable is STONE's discussion about changes over time in the transition probabilities. His assumption was that the transition probabilities follow logistic growth waves.

The first stochastic approaches to educational planning were undertaken by THONSTAD [1967] and CORREA [1963], using Markov chain models for the entire school system of a country. In the CORREA model, the transition probabilities do not represent transitions between one grade and the next. Instead, he studied transitions between different educational levels and did not differentiate between specialisations. THONSTAD's Markov model for the entire school system of a country, took a more real approach than CORREA's model; it was first published in 1967 for the Norwegian educational system which was pursued further and presented in 1969.

In spite of a considerable variation in coverage and detail, the logic behind each of these models is generally very similar, as they tend to describe the flow of students into, through, and out of the education system or some part of it. Thus, although a large number of models exist under different labels, they usually represent variations of the same principle, according to the purposes of the analysis of a particular country and to the data available.

The consequence of recognising the qualitative analysis as an important element in planning for future developments of the education system, has given emphasis in work on social indicators, questionnaire-based data collecting, field experiment, and surveys of student attitudes. A profusion of published and unpublished studies on particular parts of the educational planning process, in a variety of social settings, have been carried out since then, in order to provide a basis for the analysis to the social demand for education. 'Social demand' for education, also called 'individual demand' for education, can be defined as education as a valuable acquisition leading to a multiplicity of social goods, including employment (demand for school places by the students and their families) [see HARNQVIST, 1978]. Factors affecting demand for education have been, perhaps, the main concern underlying these studies, examples of which are the analyses developed by CONLISK [1969], MASTERS [1969], LERMAN [1974], KAYSER [1976], BOARDMAN et al. [1977], KUMAR et al. [1980], MATILLA [1982] and PSACHAROPOULOS [1982a]. Also a report published by OECD [1979] covering the experience of five countries (France, Germany, Greece, United Kingdom and Sweden) presents a detailed discussion about the main factors affecting demand, grouped in four principal categories: psychological/individual, structural/institutional, social/familial and economic/financial.

CONLISK [1969] used the Census data on children aged 5-19 in the United States to analyse the determinants of school enrolment and school performance. Dummy demographic variables describing age, colour, sex, rural-urban status, education of parents and income of parents were used to explain the changes in the dummy variables school enrolment and school performance (whether an enrolled child is behind, with, or ahead of his/her age group in years of schooling). It must be noted, however, that at this time schooling was viewed as an investment in human capital. Thus, one of the conclusions of the study is that the early stages of the investment are crucially important in determining the success of the later stages. This means that if a child starts to fall behind in schooling, he/she will tend to fall further and further behind as time passes, and eventually will tend to drop out of school sooner than the average. CONLISK also argues that an increase in parents' incomes will result in a significant increase in their children's school enrolment and performance.

Using data from a 1/1 000 sample of the United States 1960 Census, MASTERS [1969] attempted to estimate the degree of inequality of educational opportunity at the secondary level. The repetition probabilities and the drop-out probabilities for children from different family backgrounds were estimated. The results show that for children whose parents have little education or income, the repetition and drop-out probabilities are much greater than for children where both parents have graduated from high school and have high income.

LERMAN [1974] studied young people's decision to attend school, using individual, familial and area influences as explanatory variables. The study showed that, as expected, family income and the educational attainment of the family head both exerted large positive effects on school activity. Also, differences within the highest categories of family income and of school years completed by the family head did not influence significantly the high school decision, but played a substantial role in the college decision.

The rise of interest in causal modelling during the 1970s was also shown by the social scientists who started investigating the mechanisms of social mobility. BOARDMAN et al. [1977] and KAYSER [1976] are two examples of the use of this kind of education modelling. BOARDMAN applied a number of simultaneous equation models of the educational process to a sample of twelfth grade students, including, as endogenous variables: motivation, achievement, expectations, efficacy, students' perceived parents' expectations and students' perceived teachers' expectations. The analysis emphasizes that achievement is not the only output of the educational process; motivation, expectations, efficacy and belief in the ability to control the environment are independent and also important outputs. A consistent result was that efficacy and achievement are highly positively correlated.

KAYSER [1976] suggested on the other hand, that the educational aspirations throughout high school were well described by a Markov chain model. For those with high aspirations, the process appeared second-order, while for those with low aspirations the process seemed to be first-order. One conclusion of the study is that the transition probabilities did not increase with the aspirational level held. However, parental and teachers' expectations seemed to be the most important factors in promoting change in educational goals.

An attempt to examine the relationship of drop-out behaviour and re-entry flows of a local school system, to changes in local socio-economic factors, school related conditions and aspects of labour market participation was carried out by KUMAR et al. [1980]. A simulation model was developed to include these interrelated factors. The results imply that socio-economic and school environments have a stronger impact on drop-out rates than on re-entry rates. While anticipated reductions in pupil-teacher ratio and a taxable family income are likely to raise the future levels of drop-out rates. Changes in family size (a decrease), unemployment rate (an increase) and the proportion of high enrolment schools (a decrease) would tend to depress them. With regard to the magnitude of the impacts of the various factors, family size is the most influential of those under consideration (a fall in the family size would lower drop-outs and re-entry rates). A paradoxical result has, however, emerged from this study: the socio-economic and school related factors that may help improve the re-entry rates in the future are also the ones that are likely to raise the drop-out rates.

The study undertaken by MATILLA [1982] used aggregate time-series data rather than longitudinal survey data or Census data, which is usually used in the social demand approach when qualitative analysis is performed. The time-series data were organized by age and level of enrolment to study the way that male school enrolment rates respond to changes in the expected rate-of-return to schooling, the unemployment rate, the proportion of high school graduates, and variables related to participation in the armed forces. The conclusion of this study was that a small set of variables explained most of the variation in the high school and college enrolment rates of young males, the estimated rate-of-return to schooling being a strong and positively significant explanatory factor, even when other measures of family income, youth earning power, job opportunities and educational attainment were added to the model.

Under a technical cooperation program with OECD, Portugal has initiated a project aimed at understanding the determinants of the social demand for education. Social demand should be interpreted as the demand for school places by students. Questionnaires were administered to a national random sample of students in the sixth, ninth and eleventh grades in May 1979, as well as to a sample of the eleventh grade in May 1980 following a reform of subject choice in the upper secondary level of education. The results [see PSACHAROPOULOS (1982a), SOARES et al. (1980)] indicate a high degree of realism of students' expectations regarding their future economic role. Age, family income, school grades and school type were found to exert a significant influence on the decision to continue to a higher level of education, for the sixth and ninth grade samples.

The students in the eleventh grade are faced with two problems: (i) the 'numerus clausus' established for entrance to the university system; (ii) a high youth unemployment rate. A comparison between the two questionnaires administered to the eleventh grade students showed a real students' perception of the difficulties of obtaining higher education; they indicated that subject aspirations were not towards what they would "like" to study, but towards that subject for which it would be feasible to obtain admission. Here, of course, the family income emerges as an important selective factor on the preferences for post-secondary private schooling.

Although qualitative analysis became an important element in educational planning, decision-makers still are interested in quantifiable features and quantitative analysis continues to be performed to predict future enrolments. No relation exists, however, between both type of social demand approach, the Portuguese experience being no exception. It is this lack between the traditional modelling of the educational process and the more recent way of social demand approach that the present study attempts to overcome. As further described in section 2.3., the present study tries to improve the quantitative modelling of the school enrolment process. The Portuguese educational system is used to test the model, and the results and recommendations of the qualitative analyses undertaken have been taken into account in the selection of the explanatory variables which have been included in the extended model.

2.2.3 The rate-of-return approach

During the 1970s great attention was paid to the analysis of the relationship between earnings and education, with emphasis on the qualitative nature of this relationship and the relative values of the educational investment at the different levels of education. Employers' surveys have been used and earnings functions computed; the price of graduates have been analysed in order to give information on the degree of substitution between different kinds of manpower in production. Despite the conceptual and technical problems, rates-of-return (social and private) to educational investment have been calculated [see PSACHAROPOULOS (1973; 1975; 1981a; 1981b), ECKAUS et al. (1974), SOARES et al. (1982)].

The major objection to the rate-of return approach refers to the wage as a poor indicator of an individual's marginal productivity: extra earnings associated with education may be due, for example, to intelligence, to social origins, to the families' cultural backgrounds, and to factors other than education. The benefit obtained by calculating the rate-of-return on education is a sign of the present value of income derived from schooling. However, there are many other benefits, such as personal, social and external benefits, which are not measured by earnings. A limitation of this type of approach is that it is a marginal analysis, suggesting directions but not the magnitude of change. For its calculation, detailed cross-section age-income data on the current labour market is required, with one or another level or kind of schooling, but the time patterns used in this way are not historic or development time. BLAUG (1981) referred to this analysis:

2.3. The Extended Model and the Social Demand Approach

The literature shows that, since the change of attitude in the 1970s towards the premises of educational planning, the social demand approach (in a qualitative sense) and cost-benefit analyses are the approaches that have dominated the attention of researchers in educational planning up to the present. Furthermore, these analyses, in particular the social demand studies, have shown the students' educational plans, school enrolment and school performance being influenced by causal factors, such as parental, socio-economic or school environment.

The results of these studies indicate, therefore, the weakness of the existing models for educational planning, as they are based on the assumption of no interdependence between the transition rates. The transition rate projections being made did not take into account that changes in one transition rate must affect other rates, or that causal factors may also effect changes in these transition rates.

Despite all their weakness, macro analyses using flow models are still used by the planners to project the future school enrolment. Following STONE's assumption that the transition probabilities follow logistic trends, a simplified version of the logit approach to project trends on the transition rates using causal factors (generalised logit approach) was outlined by THONSTAD [1981, pp.62-63] in a statistical report published by UNESCO. However, and using THONSTAD words, 'to the best of our knowledge not much has been done in the area of simultaneous estimation of a set of transition rates' (p.62). Thus, although there could be different methods of estimating the transition probabilities, econometric estimation was the procedure selected to estimate simultaneously the parameters of the models used in this study.

The above suggests, therefore, that attempts should be made in this area. Extending the traditional Markov chain model of school enrolment by allowing the transition probabilities to be affected by changes in causal variables is the main purpose of this thesis. Linear behavioural relationships will be incorporated in the model and simultaneous estimation of time-varying transition probabilities will be performed (Chapters 6-8).

It must be noted, however, that the objective of this model is not to project future school enrolment; rather it should be seen as a more general model, with the aim of increasing understanding of the functioning of the educational system. This kind of model may be a useful supplement to the social demand approach, by indicating the directions and magnitudes of the impact of causal factors on the parameters describing school enrolment.

Chapter 3

MARKOV MODELS

3.1. Introduction

This chapter attempts to draw together the threads of the theoretical basic model in order to set the stage for subsequent applications to the Portuguese educational system. Various approaches to the study of school enrolment, such as Markov methods, regression analysis and simulation, have been used to produce projections of the number of students enrolled, graduates and drop-outs. Among the approaches, the Markov chain approach appears to be the most widely applied.

The basic concept of a Markov chain was first given by A.A. MARKOV [1906]; the mathematical formulation was developed by KOLMOGOROV [1937], DOEBLIN [1937; 1940] and DOOB [1942;1945]. The first application of a Markov model to the learning process was in psychology [MILLER, 1962]. BARTHOLOMEW [1967; 1973; 1982] has traced the use of Markov models in the social sciences.

The Markov model is a stochastic one, the theories formulated beginning with a very simple probabilistic model of some aspect of individual 'mobility'. They then invoke the central limit theorem to obtain a steady state distribution, or equilibrium distribution, that more or less approximates previous known derived distributions. A forecasting model may then been built which has no support in a theoretical framework. To be certain that a model is really valid for prediction purposes it must describe with a reasonable confidence the historical trend of observed values. However, scarcely any model applied to the education sector have been tested to determine if it generates the historical trend.

A Markov chain, or first order Markov model, is one of the simplest stochastic processes, with the following fundamental assumptions:

1. The probability of an individual moving from one state to another depends only on the states occupied and not on the past history of transitions.

2. Each state is assumed to be homogeneous with respect to the transition probabilities. This means that every individual who finds himself in a particular state is assumed to have the same probability of moving to another state.

3. Each state is assumed to be independent of the others as far as their transition probabilities are concerned.

All the approaches used to study the behaviour of school enrolment, using Markov processes, assume that the educational system can be described by a stationary Markov chain; that is, the transition probabilities are constant over the sampling periods. Stochastic models applied to school enrolment have, therefore, been entirely passive in terms of behaviour. The state of the system at time t depends only on the state of the system at time t-1; past academic performance has no significant influence on the probability of a student being promoted or repeating. The past can be incorporated by extending the idea of the present to include some of the past movements to give 2nd, 3rd and higher order chains. Nevertheless, students' behaviour may not be time invariant, especially when their expectations and motivations, for instance, may be disturbed by exogenous factors. The number of students enrolled at a certain educational level can be assumed to be a function of certain variables (some of them difficult to measure, as they express qualitative features) such as economic and sociological characteristics of the population, direct costs of acquiring a certain level of education, the earnings of graduates of different levels, the opportunities of getting good employment, etc. Therefore the transition probabilities may change through time and a non-stationary Markov chain may result.

The main difficulty which arises when trying to apply a Markov model to a social system is the estimation of the transition probabilities. Dealing with human beings, the most satisfactory and desirable procedure is to estimate these transition probabilities from individual observations. This, however, involves the existence of a statistical data base unavailable in most situations. Usually only the aggregate proportions relating the number of students in each grade for each time t are known.

A stationary Markov chain can be a useful first step, a valid approach to support the formulation of a new mathematical model which includes some theoretical reasoning. The following sections of this chapter will consist of a description of the stationary Markov chain applied to school enrolment, as well as of the different processes of estimating the transition probabilities.

3.2. The Concept of a Markov Model

Markov models are particularly useful for describing and analysing the nature of changes generated by the movement from one state to another. A Markov model is a mathematical model of the "behaviour" (in a loose sense) of individual units where movements from one state to another are determined by a probabilistic rule. The traditional Markov model is what is called a stochastic process, which means that the model attaches probabilities to the various possible states of the system and the process develops according to these probabilities. The Markov chain, or first-order Markov model, is the simplest one as it relates the state of a system at time t with the state of a system at time t-1; it is independent of the history of the system prior to t-1. In a second order Markov model the state of the system at time t would depend on the states of the system at both time t-1 and t-2.

The information relating to the observed probabilities of past trends can be organised into a matrix which is the basic framework of a Markov model. Let the states of the Markov chain be numbered 1, 2, ... ,s. Denoting by p_{ij} the transition probability from state i to state j, i =1,2, ... ,s and j =1,2, ... ,s, the matrix of these transition probabilities can be illustrated as follows:

$$\underline{P} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & \cdots & p_{1s} \\ p_{21} & p_{22} & p_{23} & \cdots & p_{2s} \\ p_{31} & p_{32} & p_{33} & \cdots & p_{3s} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ p_{s1} & p_{s2} & p_{s3} & \cdots & p_{ss} \end{bmatrix}$$

These satisfy the condition that $\sum_j p_{ij} = 1$, that is, the row-sums are equal to unity.

Assuming the system is stationary, that is each state of the system is homogeneous, the transition probability matrix remains constant over time; taking $\underline{n}'(0)$ to be a row vector of the initial state (at time 0), the state of the system at time 1 can be obtained by multiplying the initial state vector $\underline{n}'(0)$ by the transition probability matrix $\underline{P}$.

$$\underline{n}'(1) = \underline{n}'(0) \cdot \underline{P}$$

the state of the system at time 2 can be obtained by:

$$\underline{n}'(2) = \underline{n}'(1) \cdot \underline{P} = \underline{n}'(0) \cdot \underline{P}^2$$

and in the general case:

$$\underline{n}'(t) = \underline{n}'(0) \cdot \underline{P}^t$$

or

$$\underline{n}'(t) = \underline{n}'(t-a) \cdot \underline{P}^a$$

The distribution of any variable at time t is dependent on its distribution at the initial state and the transition probability matrix $\underline{P}$. If the process follows a Markov chain then the distribution at any time in the future can be found by repeated multiplication of the vector of the initial state by the transition probability matrix.

The general theory of Markov chains shows that when t tends to infinity, the limits of $\underline{n}(t)$ and of the matrix $\underline{P}^t$ depend on the structure of $\underline{P}$ alone. Provided $\underline{P}$ is a transition matrix for a 'regular' Markov chain (i.e. all the elements of $\underline{P}(t)$ are positive entries) then the probabilities all approach limits as t tends to infinity. Two important theorems relating to the equilibrium properties are provided by KEMENY and SNELL [1967, pp.96-98]:

<u>Theorem 1</u> : If $\underline{P}$ is a transition matrix for a regular Markov chain then:

(i) $\lim_{n\to\infty} \underline{P}^n = \underline{A}$

(ii) each row of $\underline{A}$ is the same probability vector $\underline{b}'$.

(iii) the elements of $\underline{b}'$ are all positive.

<u>Theorem 2</u> : If $\underline{P}$ is a transition matrix for a regular Markov chain and $\underline{A}$ and $\underline{b}'$ are as in Theorem 1, then the unique vector $\underline{b}'$ is the unique probability vector such that $\underline{b}'\ \underline{P} = \underline{b}'$

The matrix $\underline{A}$ is defined as the 'limiting matrix'. Thus if a social process approximates to a regular Markov chain, it will approach or may already have reached an equilibrium where the proportions in each state remain constant for all future time periods. Equation $\underline{b}'\ \underline{P} = \underline{b}'$ can be solved via a set of simultaneous equations or by structural analysis using eigenvalues and eigenvectors to give a solution that does not depend on the initial state of the system. The limiting matrix will be of the form:

$$\underline{A} = \begin{bmatrix} p_1 & p_2 & \cdots & p_s \\ p_1 & p_2 & \cdots & p_s \\ p_1 & p_2 & \cdots & p_s \\ \cdots & \cdots & \cdots & \cdots \\ p_1 & p_2 & \cdots & p_s \end{bmatrix}$$

and the probability vector $\underline{b}' = (p_1\ p_2\ \cdots\ p_s)$ holds for the system in equilibrium. This result can also be derived using the Frobenius theorem. The row-sum condition on the matrix $\underline{P}$ ensures that the dominant eigenvalue is unity. Thus, the eigenvector equation becomes

b' P = b', as above [see TAKAYAMA (1974; pp.367-379)].

Some of the social systems, like the educational system, change their internal structure over time. In every time period there are students entering the system and students leaving the system. A system with these characteristics is called an open system. Such systems are studied by introducing a new state (the absorbing state) to include all the losses. The stochastic augmented transition probability matrix is thus of the form:

$$\left[\begin{array}{c|c} \underline{P} & \underline{p}_{s+1} \\ \hline \underline{0}' & 1 \end{array}\right]$$

Furthermore the transitions within the educational system are to the same or a higher grade only. When the grades are organised in increasing order, the matrix $\underline{P}$ is an upper triangular matrix and so the augmented matrix also has this form. The eigenvalues of such a matrix are the diagonal elements and so the spectral representation technique has some advantages in finding the equilibrium distribution.

In Markov chain analysis, therefore, for modeling purposes the equilibrium distribution is of interest not as a forecast of the future state of the system but as a projection of what it would be if the observed pattern of movement remains constant. The assumption of stationarity requires that the parameters remain constant throughout the predictive period and is therefore a severe assumption; an assumption that lends support to the notion that Markov chains should be used primarily for descriptive rather than predictive purposes [BROWN, 1970]. Moreover, when dealing with human beings, the assumption of stationarity seems to be a very weak assumption as human behaviour does not remain the same over time. The effects of non-stationarity is that instead of there being one transition matrix, there are t transition matrices, one for each time h (h=1, ... ,t). In this case, the state of the system at time 1 is given by:

$$\underline{n}'(1) = \underline{n}'(0) \cdot \underline{P}(1)$$

the state of the system at time 2 can be obtained by

$$\underline{n}'(2)= \underline{n}'(1) \cdot \underline{P}(2)= \underline{n}'(0) \cdot \underline{P}(1) \cdot \underline{P}(2)$$

and in the general case

$$\underline{n}'(t) = \underline{n}'(t-1) \cdot \underline{P}(t) = \underline{n}'(0) \cdot \prod_{h=1}^{t} \underline{P}(h)$$

We now move on to consider the basic Markov model.n

3.3. The Basic Markov Model

The basic equation of the Markov model for the students' mobility within the education system can be written in the form [see BARTHOLOMEW, 1973]:

$$\underline{n}'(t) = \underline{n}'(t-1)\ \underline{P}^{t}_{t-1} + \underline{n}'_{o}(t) \qquad (3.1)$$

with

$$\sum_j p_{ij} < 1 \text{ and } p_{ij} > 0 \text{ for all i and j}$$

where

$\underline{n}'(t)$ is the (1 x s) row vector whose elements are the number of students in different grades at time t, corresponding to the academic year t/t+1

$\underline{P}^{t}_{t-1}$ is the (s x s) matrix whose elements are the transition probabilities p_{ij} (probability that a student in grade i at time t-1 moves to grade j at time t)

$\underline{n}_o'(t)$ is the (1 x s) row vector of the new entrants to the education system by grade at time t

This equation gives the flow of students within the education system, relating the number of students at time t-1 with the number of students at time t through a 'flow data' matrix.

The matrix $\underline{P}_{t-1}^{t}$ is an upper triangular matrix in which the elements of the diagonal are the repetition probabilities and those in the upper triangle are the different promotion probabilities between grades.

However, equation (3.1) does not give the number of students that leave the education system at each time period, as graduates or drop-outs. It follows from the definitions of the transition probabilities that the sum of the promotion, repetition, drop-out and graduation probabilities equals unity. Assuming that s+1 is the state of the system whose elements are the graduates and drop-outs (those who leave the educational system), the following vector can be written:

$$\underline{p}_{s+1}(t) = \begin{bmatrix} 1 - \sum_j p_{1j} \\ \cdot \\ \cdot \\ \cdot \\ 1 - \sum_j p_{sj} \end{bmatrix}$$

where the elements of the vector represent the drop-out probabilities in the different grades at time t.

In order to include the drop-out probabilities in equation (3.1) the augmented matrix $\underline{P}_{a,t-1}^{t}$ defined by BARTHOLOMEW [1973, p.66] is used:

$$\underline{P}^t_{a,t-1} = \left[\begin{array}{c|c} \underline{P}^t_{t-1} & \underline{p}_{s+1}(t) \\ \hline \underline{0}' & 1 \end{array}\right]$$

This is a matrix with the appropriate form for a Markov process with one absorbing state.[1] The equation of the model can then be written as:[2]

$$\underline{n}'_a(t) = \underline{n}'(t-1)\ \underline{P}^t_{a,t-1} + \underline{n}'_{o,a}(t) \tag{3.2}$$

with

$$\sum_j p_{ij} = 1 \quad i=1, \ldots ,s+1 \text{ (row-sum condition)} \tag{3.3}$$

$$p_{ij} \geq 0 \qquad \text{(non-negativity condition)} \tag{3.4}$$

The state s+1 is the absorbing state of the system which aggregates the drop-outs and the graduates who leave the system in each time period; $\underline{n}'_a(t)$ and $n'_{o,a}(t)$ are now row vectors of size [1 x (s+1)]. Vector $\underline{n}'_a(t)$ includes as well as the number of students by grade at time t, the global number of drop-outs plus graduates that have left the school system at year t (the end of academic year t-1/t).

Assuming that all transition matrices remain constant over time, equation (3.2) can be taken for all historical years in order to write an aggregate form. Being interested in establishing the drop-outs by grade rather than the global drop-outs accumulated, the aggregate form can be written then:

$$\begin{bmatrix} n_1(1) & \dots & n_{s+1}(1) \\ \vdots & & \vdots \\ n_1(k) & \dots & n_{s+1}(k) \end{bmatrix} = \begin{bmatrix} n_1(0) & \dots & n_s(0) \\ \vdots & & \vdots \\ n_1(k-1) & \dots & n_s(k-1) \end{bmatrix} \left[\begin{array}{c|c} \underline{P}^t_{t-1} & \underline{p}_{s+1} \end{array} \right]$$

$$+ \begin{bmatrix} n_{o1}(1) & \dots & n_{os}(1) & 0 \\ \vdots & & \vdots & \vdots \\ n_{o1}(k) & \dots & n_{os}(k) & 0 \end{bmatrix} + \begin{bmatrix} u_1(1) & \dots & u_{s+1}(1) \\ \vdots & & \vdots \\ u_1(k) & \dots & u_{s+1}(k) \end{bmatrix}$$

with $t = 1, \dots, k$ and where $\underline{u}_i(t)$ are $[1 \times (s+1)]$ row vectors of disturbance terms.

Each column of matrix $\underline{A}$ can be written as

$$\underline{n}_i = \begin{bmatrix} n_i(1) \\ \vdots \\ n_i(k) \end{bmatrix} = \begin{bmatrix} n_1(0) & \dots & n_s(0) \\ \vdots & & \vdots \\ n_1(k-1) & \dots & n_s(k-1) \end{bmatrix} \begin{bmatrix} p_{1i} \\ \vdots \\ p_{si} \end{bmatrix} + \begin{bmatrix} n_{oi}(1) \\ \vdots \\ n_{oi}(k) \end{bmatrix} + \begin{bmatrix} u_i(1) \\ \vdots \\ u_i(k) \end{bmatrix}$$

for all $i=1, \dots, s+1$

or

$$\underline{n}_i = \underline{N}_i \, \underline{p}_i + \underline{n}_{oi} + \underline{u}_i \qquad i=1, \dots, s+1 \qquad (3.5)$$

This expression stacks over time the values corresponding to grade i. More precisely, this expression relates the number of students in grade i at time t to the number of students in all grades at time t-1, through a vector of the different probabilities of moving into grade i.

Stacking these values not only over time but also over grades, the historical observed values can then be aggregated in the form:

$$\begin{bmatrix} \underline{n}_1 \\ \cdot \\ \cdot \\ \cdot \\ \underline{n}_s \\ \underline{n}_{s+1} \end{bmatrix} = \begin{bmatrix} \underline{n}_{o1} \\ \cdot \\ \cdot \\ \cdot \\ \underline{n}_{os} \\ \underline{0} \end{bmatrix} + \begin{bmatrix} \underline{N}_1 & \cdots & \underline{0} & \underline{0} \\ & \ddots & & \vdots \\ \underline{0} & \cdots & \underline{N}_s & \underline{0} \\ \underline{0} & \cdots & \underline{0} & \underline{N}_{s+1} \end{bmatrix} \begin{bmatrix} \underline{p}_1 \\ \cdot \\ \cdot \\ \cdot \\ \underline{p}_s \\ \underline{p}_{s+1} \end{bmatrix} + \begin{bmatrix} \underline{u}_1 \\ \cdot \\ \cdot \\ \cdot \\ \underline{us} \\ \underline{u}_{s+1} \end{bmatrix}$$

or

$$\underline{n} = \underline{n}_o + \underline{N}\ \underline{p} + \underline{u} \qquad (3.6)$$

where $\underline{n}$, $\underline{n}_o$ and $\underline{u}$ are [k(s+1) x 1] column vectors, $\underline{p}$ is a [s(s+1) x 1] column vector and $\underline{N}$ is a [k(s+1) x s(s+1)] block diagonal matrix with $\underline{N}_1 = \underline{N}_2 = \ldots = \underline{N}_{s+1}$. Equation (3.6) can be rewritten in the following form:

$$\underline{n}^* = \underline{N}\ \underline{p} + \underline{u} \qquad (3.7)$$

with $\underline{n}^* = \underline{n} - \underline{n}_o$

The stochastic assumption for the disturbance term is

$$E(\underline{u}) = \underline{0}$$

and (3.8)

$$E(\underline{u}\,\underline{u}') = \underline{\Sigma}$$

where $\underline{\Sigma}$ is the covariance matrix of type [k(s+1) x k(s+1)]

3.4. Estimating the Transition Probabilities of the Model

LEE, JUDGE and ZELLNER [1970] have undertaken several experiments using different estimating procedures applied to individual data and aggregate data, to estimate the transition probabilities of a stationary Markov model. The results of these experiments have shown that if micro data are available, the maximum likelihood estimator applied to individual data is superior to any other estimator using aggregate data. The restricted least squares estimators are better than the unrestricted least squares estimators, the restricted generalised least squares (GLS) (or maximum likelihood (ML) or minimum chi-square (MCS)) being, therefore, the recommended method.[3]

Because of the unavailability of individual data, the study will involve the problem of estimating the transition probabilities from aggregate data.

The use of restricted estimators ensures that estimates will fall in the admissible region of the parameter space. But although the unrestricted ordinary least squares (OLS) estimator is unbiased, when the restrictions occur a quadratic programming procedure is used to derive the estimates. Also, if more than one independent variable is involved, it is difficult to evaluate the moments and obtain the sampling properties of the restricted estimator [ZELLNER, 1961]. Moreover, heteroscedasticity may be present; that is, the variances of the disturbance terms may not be constant and the variance-covariance matrix may not have the form of a positive scalar times an identity matrix. The estimates are then inefficient and the estimated covariance matrix is biased and inconsistent. In an attempt to correct the problem of heteroscedasticity AITKEN's [1934] generalised least squares method might be used.

3.4.1. The OLS Estimator

It is assumed that the number of students by grade at time t is generated in a consistent way with a first order stationary Markov chain. Proceeding in the usual manner, each observed value $n_i(t)$ $(t = 1, \ldots, k)$ has associated with it some random disturbance $u_i(t)$, which may be represented, in general, in an aggregated form, by the expression (3.7) presented in the previous section:

$$\underline{n}^* = \underline{N}\,\underline{p} + \underline{u}$$

with

$$E(\underline{u}) = \underline{0}$$

$$E(\underline{u}\,\underline{u}') = \underline{\Sigma}$$

where $\underline{\Sigma}$, the variance - covariance matrix, is a $[k(s+1) \times k(s+1)]$ non-diagonal matrix.

The least squares estimator (OLS) of the transition probability vector is given by:

$$\underline{\hat{p}} = (\underline{N}'\underline{N})^{-1}\ \underline{N}'\underline{n}^{*} \qquad (3.9)$$

where $\underline{N}'\underline{N}$ is a non-singular block diagonal matrix with matrices $\underline{N}_i'\underline{N}_i$ on the main diagonal.

This estimator, however, does not guarantee that the non-negativity and row-sum conditions [equations (3.3) and (3.4)] are satisfied. The row-sum codition can be easily met by adding the restrictions and setting up the Lagrangean. The same does not happen with the non-negativity condition which may be violated by the unrestricted OLS estimator. Generating estimates that lie outside the unit interval can be avoided by imposing constraints.

The problem then becomes that of finding the estimate $\underline{\hat{p}}^{s}$ which minimises the positive quadratic form:

$$\Phi = \underline{u}'\underline{u} = (\underline{n}^{*} - \underline{N}\ \underline{\hat{p}})'(\underline{n}^{*} - \underline{N}\ \underline{\hat{p}}) \qquad (3.10)$$

subject to constraints

$$\begin{aligned} \underline{G}\ \underline{\hat{p}} &= \underline{\eta}_{s+1} \\ \underline{\hat{p}} &> \underline{0} \end{aligned} \qquad (3.11)$$

where $\underline{G}$ is a $[(s+1) \times (s+1)]$ known coefficient matrix $[\underline{A}_1 \ldots \underline{A}_{s+1}]$ with each $\underline{A}_i$ a $[(s+1) \times (s+1)]$ diagonal matrix of entries zero or unity on the main diagonal, and $\underline{\eta}_{s+1}$ is an $[(s+1) \times 1]$ column vector with all entries equal to unity. The first constraint $\underline{G}\ \underline{\hat{p}} = \underline{\eta}_{s+1}$ is a different way of writing the relationship between the transition probabilities given by equation (3.3).

This is a quadratic programming problem as restrictions are linear and the objective function is a quadratic form. Using the KUHN-TUCKER [1951] equivalence theorem for non-linear programming and the duality theorem of DORN [1960] for quadratic programming, the problem can be reduced to the primal-dual linear programming problem (see A.1, Appendix A).

The problem is then

- find (λ_1, λ_2, $\underline{\hat{p}}^s$) that maximises

$$- (\underline{\lambda}_1' \underline{\alpha}_1 + \underline{\lambda}_2' \underline{\alpha}_2 + \underline{\beta}' \underline{\hat{p}}) \qquad (3.12)$$

subject to

$$\begin{aligned}
&\underline{G}\,\underline{\hat{p}} + \underline{\alpha}_1 = \underline{n}_{s+1} \\
&-\underline{G}\,\underline{\hat{p}} + \underline{\alpha}_2 = -\underline{n}_{s+1} \\
&\underline{G}'\,\underline{\lambda}_1 - \underline{G}'\,\underline{\lambda}_2 = \underline{N}\,\underline{n}^* - \underline{N}'\underline{N}\,\underline{\hat{p}} + \underline{\beta} \\
&\underline{\hat{p}}\ ,\ \underline{\lambda}_1,\ \underline{\lambda}_2,\ \underline{\alpha}_1,\ \underline{\alpha}_2,\ \underline{\beta} \quad \geq \underline{0} \qquad (3.13)
\end{aligned}$$

where $\underline{\lambda}_1$ and $\underline{\lambda}_2$ are the [(s+1) x 1)] vectors of dual variables and $\underline{\alpha}_1$, $\underline{\alpha}_2$, $\underline{\beta}$ are the slack vectors. This problem can be solved by using the standard simplex version algorithm developed by WOLFE [1959].

3.4.2. The GLS Estimator

The estimated vector $\underline{\hat{p}}$, obtained by solving the previous linear programming problem, ensures that the non-negativity and row-sum conditions are satisfied. However, the specification of the variance-covariance matrix $\underline{\Sigma}$ has been ignored. The variances of the disturbance terms may not be constant and the covariance matrix does not have the proper form $E(\underline{u}\ \underline{u}') = \sigma^2\ \underline{I}$; that is, it is not of the form of a positive scalar times an identity matrix. Heteroscedasticity would be expected to be present so the OLS estimator, even restricted, is not efficient. The estimated covariance matrix is biased and inconsistent so that the standard tests for significance do not apply. To overcome such heteroscedasticity it is necessary to weight the original data and then perform OLS estimation upon the transformed model; that is, use generalised least squares (GLS). If s of the equations of the model are known, the remaining equation is therefore determined by the row-sum condition. One of the equations of the model may then be deleted and the reduced form written as:

$$\begin{bmatrix} \underline{n}_1^* \\ \cdot \\ \cdot \\ \cdot \\ \underline{n}_s^* \end{bmatrix} = \begin{bmatrix} \underline{N}_1 & & \underline{0} \\ & \ddots & \\ \underline{0} & & \underline{N}_s \end{bmatrix} \begin{bmatrix} \underline{p}_1 \\ \cdot \\ \cdot \\ \cdot \\ \underline{p}_s \end{bmatrix} + \begin{bmatrix} \underline{u}_1 \\ \cdot \\ \cdot \\ \cdot \\ \underline{u}_s \end{bmatrix}$$

or

$$\underline{n}^* = \underline{N}\ \underline{p} + \underline{u} \qquad (3.14)$$

with

$$E(\underline{u}) = \underline{0}$$

$$E(\underline{u}\ \underline{u}') = \underline{\Sigma} \qquad (3.15)$$

where $\underline{\Sigma}$ is a (sk x sk) non singular matrix. Using $\underline{\Sigma}$ as a weighted matrix in forming the weighted least squares, the unrestricted GLS estimator is given by

$$\hat{\underline{p}} = (\underline{N}'\ \underline{\Sigma}^{-1}\ \underline{N})^{-1}\ (\underline{N}'\ \underline{\Sigma}^{-1}\ \underline{n}^*) \qquad (3.16)$$

The deleted parameter $\underline{p}_{s+1}$ may be estimated using the expression

$$\hat{\underline{p}}_{s+1} = \underline{n}_{s+1} - \underline{R}\ \hat{\underline{p}} \qquad (3.17)$$

where $\underline{R}$ is a submatrix of $\underline{G}$, with the form $[\underline{A}_1, \ldots, \underline{A}_s]$ with each $\underline{A}_i$ a $[(s+1) \times (s+1)]$ diagonal matrix with entries zero or unity.

Taking into account the restrictions and including them in the model, the estimation problem for getting the restricted GLS estimators becomes again a quadratic programming problem that can be solved by using the simplex algorithm. The primal-dual formulation is deduced in A.3 (see Appendix A) with the final form:

- find $(\ \underline{\lambda}, \hat{\underline{p}}^s)$ which maximises

$$-\ (\ \underline{\lambda}'\ \hat{\underline{p}}_{s+1} + \underline{\beta}'\ \hat{\underline{p}}\ ) < 0 \qquad (3.18)$$

subject to

$$\begin{aligned} &\underline{R}\ \hat{\underline{p}} + \hat{\underline{p}}_{s+1} = \underline{n}_{s+1} \\ &\underline{R}'\ \underline{\lambda} + \underline{N}'\ \underline{\Sigma}\ \underline{N}\ \hat{\underline{p}} - \underline{\beta} = \underline{N}'\ \underline{\Sigma}^{-1}\ \underline{n}^* \qquad (3.19) \\ &\hat{\underline{p}}, \hat{\underline{p}}_{s+1}, \underline{\lambda}, \underline{\beta} > \underline{0} \end{aligned}$$

Chapter 3

Footnotes

1. As already mentioned, an absorbing state is one which once entered cannot be left, that is, there is a zero probability of leaving it. From every non-absorbing state it is possible to go to one of the absorbing states of the system.

2. This equation gives the global drop-outs accumulated. The drop-outs by grade can be obtained using

$$\underline{n}'_d(t) = \underline{n}'(t-1)\, \underline{Q}_t$$

where $\underline{n}'_d(t)$ is the (1 x s) row vector of the number of students who leave the system at time t and $\underline{Q}_t$ is the diagonal matrix of size (s x s) whose non-zero elements are those of vector $\underline{p}_{s+1}(t)$.

3. LEE, JUDGE and ZELNER [1970] have proved that the expressions for the estimators of transition probabilities obtained by GLS and MCS are the same. Also, the minimum chi-square estimator (MCS) is obtained by minimising the chi-square error while the maximum likelihood (ML) estimator is obtained by maximising the likelihood function. These approaches may be considered as corresponding to primal and dual problems, so the three methods are equivalent.

Chapter 4

APPLICATION OF THE BASIC MARKOV MODEL TO THE PORTUGUESE EDUCATIONAL SYSTEM

4.1. Introduction

This chapter will concentrate on the application of the basic Markov model, described in Chapter 3, to a subsystem of the Portuguese educational system. In Portugal, as in many European countries, at the end of each academic year, a student needs to reach a certain level of attainment in order to progress to the next grade; otherwise the student fails and either repeats the grade or leaves the system and is considered as a drop-out. Thus, at the end of each academic year the student is faced with three possible situations: (i) being promoted to the next grade and carrying on his studies; (ii) failing the grade and staying in the same grade as the previous year; (iii) deciding to leave the school system, having succeeded or not in his/her studies. It is this kind of anual grading system that makes the Markov model appropriate to describe the movements of the students within the Portuguese educational system.

The analysis undertaken in this chapter applies to the whole country;[1] the least squares estimation procedure will be used to produce estimators for the different transition probabilities. Assuming that certain disturbances in the data are due to the return of students from the old colonies Angola and Mozambique after the revolution that took place in April 1974, an iterative process is applied in order to separate these students from the observed data and to obtain a "corrected data matrix". The least squares estimation procedure will be used to calculate new estimators for the transition probabilities.[2]

4.2. The School System

The main structure of schooling of the Portuguese educational system is set out in Table 4.1. Until the middle of the 1970s, kindergartens were provided only by the private sector. The government's intention of gradually expanding kindergartens is now, however, limited by the availability of financial resources needed for the provision of trained staff and adequate equipment. At present, compulsory schooling officially consists of the completion of the basic level of education, which involves the four primary and the two preparatory grades, or the attainment of the age of fourteen. However, an extension of the basic schooling to nine years, which includes the three unified grades of the secondary level or to age fifteen, which ever occurs first, seems likely to be introduced within a few years. The higher education sector comprises a university subsystem (providing a four or five years course leading to a 'licenciatura') and a non-university subsystem (a three year course leading to a bachelor's degree).

An expansion of school attendance since the 1960s has generated a consequent increase of the educational attainments of the Portuguese population. Compulsory schooling for girls was not introduced until 1960 and even then was only for four years, until it was extended to six years in 1964. The distribution of the Portuguese population over the age of fourteen, by educational

Table 4.1

The Portuguese School System Structure

Level	Type	Grade	Age
Kindergarten	Pre-primary		3-6
Basic	Primary	1-4	6-10
	Preparatory	5-6	10-12
Secondary	Unified General course	7-9	12-15
	Complementary course	10-12	15-18
Higher	University	13-17	≥18
	Non-University*	13-15	>18

*includes Polytechnics

Table 4.2

Percentage Distribution of the Population by Level of Educational Attainment

Level of educational attainment	Year 1970	1981
Illiterate	28.2	19.3
Without Primary education	22.3	12.7
Primary education	35.9	41.8
Preparatory education	9.5	12.8
Secondary education	3.2	7.8
Medium education	0.1	0.6
Higher education	0.8	1.4
TOTAL	100.0	100.0

SOURCE: National Institute of Statistics (INE), Census 1970 and 1981 (include the Islands of Madeira and Azores), Portugal,1972; 1983.

attainments, is compared for 1970 and 1981 in Table 4.2. The percentage of the population with education beyond the primary level has doubled between 1970 and 1981. However, the bulk of the population has still at most a primary education; 86.4% in 1970 and 73.8% in 1981. The rate of adult illiteracy was still high in 1981, despite the fact that all young people should now become literate through attendance at school. Nevertheless, the average age of illiterates is rising, which implies a significant decrease of illiteracy rates over the next few years.

In 1981 students comprised about 20% of Portugal's population of about 10 million, as compared to 16% in 1970. Of the total student population, 46% were in the basic primary level, 18% in the basic preparatory level, 12.5% in the secondary unified general course, 8.5% in the secondary complementary course and 3.7% in the higher education level. The distribution of students in 1981 is compared to that of 1970 in Table 4.3. In 1970 the bulk of the student population received only four years of schooling. The percentage of students who continued their studies beyond grade 4 has markedly increased between 1970 and 1981. Only 10.4% of students continued their schooling beyond this level in 1970 compared with 18% in 1981. This expansion of enrolment in the preparatory level resulted from the scheme that increased compulsory education from 4 to 6 years in 1964. However, the first generation of children to benefit moved into the preparatory schools only in 1968. It was, therefore, during the 1970s that preparatory schooling had its biggest expansion. The availability of this stage of schooling has led to an increased demand for the next stage. Only 3% of school leavers were admitted to a university (10% of the students in higher education are enrolled in non-university courses).

Table 4.3

Percentage Distribution of Students by Level of Education

Level of education	Year 1970	1981
Primary education	65.0	46.0
Preparatory education	10.4	18.0
Secondary-unified general course	15.7	12.5
Secondary-complementary course	3.1	8.5
Higher education	3.3	3.7
Others*	1.5	12.3
TOTAL	100.0	100.0

*include Pre-primary schooling and medium education (teachers training, nursing, artistic courses and others)

SOURCE: Statistics of Education, 1970; 1981. INE, Portugal, 1971; 1983.

4.3. The Scope of the Analysis

This study will concentrate on the basic preparatory and secondary levels of education (grades 5-11). The reasons for disregarding the primary and university levels are different, but both are sufficient by themselves to exclude these levels from the scope of the analysis.

The structure of primary education has changed from the school-year 1975/76. Until then, primary education contained four grades in each of which a student could be (i) promoted to the next grade, (ii) fail and stay in the grade, or (iii) leave the school system. From the school-year 1975/76 the primary system has included two phases. A student entering the first phase stays two years in that phase before being transfered, or not, to the second phase, in which he/she has to stay another two years. The flow tables between years for primary education require different coefficients, as a "retention" within the phase results for those students who have been only one year in the phase. The analysis for primary education would therefore need to be taken separately from the other levels. This is why this level has been excluded from the present study.

The enforcement of a 'numerus clausus'[3] on University entrance, established every year by the academic institutions according to their capacities, breaks the link, in Markov model terms, between the second and third levels of education, as the flow between these two levels cannot be studied as a stochastic process. Furthermore, the 'financial autonomy' given to the universities leads to the employment of resource allocation models being used by the institutions. The analysis by course is essential in this case.

The twelfth grade is also excluded from this analysis as it is a very recent school grade, introduced at the end of the secondary level, for which there are not yet enough available data. This 12th grade was introduced in 1980/81 and fills a year between the previous end of secondary schooling (11th grade) and the start of higher education.

4.4. The Data and the Variables of the Model

The data used in the estimation of the model refers to the public sector schools and covers the last available school years (1970/71 - 1981/82) for the seven grades included in the analysis. One of the main concerns while working with a model is the data gathering and its coherence. However, the data characteristics raise a question relative to the impact of the choice of the length of the time-series sample period adopted. The unavailability, since 1978/79, of the official Statistics of Education, usually published by the National Institute of Statistics (INE), has caused a break in the continuity between these data and those published by the government departments, as they are gathered in different periods of the year. Also, the enlargement of the time-series to the 1960s would bring the problem of overlap with changes in the structure of the educational system and consequently would cause serious inadequacies in the time-series data. To avoid these problems the study uses the data recently published by the Educational Planning Bureau (GEP) of the Ministry of Education of Portugal, which covers twelve years (1970/71 - 1981/82). The data used in the present study are presented in Table 4.4. Grades 5 and 6 refer to the preparatory level and grades 7 to 11 refer to the secondary level. Table 4.5 shows the structure of the transition matrix $\underline{P}^{t}_{a,t-1}$, defined in Chapter 3, for this case study.

Table 4.4

Enrolment by Grade in the Preparatory and Secondary Education - Whole Country*

Time	Year	Preparatory level			Secondary level			
		Grade 5	Grade 6	Grade 7	Grade 8	Grade 9	Grade 10	Grade 11
0	1970/71	92599	64315	48158	38202	21654	13087	8466
1	1971/72	99769	73647	52015	42239	26499	13051	10519
2	1972/73	115137	83987	60249	44534	33340	16875	10983
3	1973/74	121980	98906	81206	47341	36798	21110	14445
4	1974/75	117124	105879	98785	69909	48705	41156	20129
5	1975/76	145407	115997	104322	77394	62547	37196	37693
6	1976/77	139809	127153	92023	108500	77348	45976	44658
7	1977/78	143749	120590	101845	88998	104730	44029	52426
8	1978/79	151456	121383	91417	73778	72335	43770	38354
9	1979/80	152338	124343	94553	70132	72704	41554	39534
10	1980/81	168327	129714	97738	74893	63803	46246	47867
11	1981/82	178243	137784	101797	77250	64405	47953	57119

* These data do not include the islands of Madeira and Azores, excluded from this study.

SOURCE: Diagnostico/Previsoes, Educational Planning Bureau (GEP), Ministry of Education, 1983.

Table 4.5

The Augmented Transition Matrix $\underline{P}^{t}_{a,\ t-1}$ for the Case Study

Grades (Year t)	Grades (Year t+1)							
	5	6	7	8	9	10	11	Drop-outs (Year t)
5	p_{55}	p_{56}						$p_{5,\ s+1}$
6		p_{66}	p_{67}					$p_{6,\ s+1}$
7			p_{77}	p_{78}				$p_{7,\ s+1}$
8				p_{88}	p_{89}			$p_{8,\ s+1}$
9					p_{99}	$p_{9,10}$		$p_{9,\ s+1}$
10						$p_{10,10}$	$p_{10,11}$	$p_{10,\ s+1}$
11							$p_{11,11}$	$p_{11,\ s+1}$*
								1

* this value covers all students who finished the 11th grade as well as those who have failed this grade and have left the public school system.

NOTE: here and in the following tables a blank space means zero.

The observed point estimates of the transition probabilities are shown in Table 4.6. These point estimates have been derived from a breakdown of enrolment by grade into those entering the grade for the first time, and those repeating the grade. These point estimates are themselves not fully satisfactory and an examination of Table 4.6 shows that both non-negativity and row-sum conditions are violated. However, the existence of these point estimates allows a valuable point of reference for subsequent Markov modelling of promotions and repetitions in the Portuguese educational system. Throughout the rest of this thesis, the outcomes of the models used will be continually compared with these point estimates.

A simple look through the data in Table 4.6 shows these probabilities are non-stationary and the non-negativity condition together with the row-sum condition are violated for some of the observed point estimates. Although the time-series used in this study is consistent in terms of schooling structure and data gathering, the period of analysis presents big disturbances. It is reasonable to assume that these disturbances are due to the return of people from the two old colonies, Angola and Mozambique, after the revolution that took place in April 1974. The number of student returnees who were received by the public school system during the second half of the seventies is unknown. However, it is known that about 700 000 people had to be absorbed into the community, half of these being less than sixteen years old. Table 4.7. shows that the school enrolment increased from 1970 to 1978 in the same way as from 1960 to 1970. During the period 1960 to 1981 there was an overall increase of more than two-thirds in school enrolment. However, most

Table 4.6 The Observed Point Estimates of the Transition Probabilities - Whole Country

Trans Probab.	p_{55}	p_{56}	p_{66}	p_{67}	p_{77}	p_{78}	p_{88}	p_{89}	p_{99}	$p_{9,10}$	$p_{10,10}$	$p_{10,11}$	$p_{11,11}$	p_{5d}	p_{6d}	p_{7d}	p_{8d}	p_{9d}	$p_{10,d}$	$p_{11,d}^{*}$
<u>YEAR</u>																				
1971	.16	.67	.17	.63	.23	.67	.26	.58	.21	.58	.04	.77	.11	.17	.19	.10	.16	.21	.23	.89
1972	.15	.73	.16	.66	.22	.68	.22	.64	.23	.64	.06	.74	.12	.12	.18	.10	.14	.12	.20	.88
1973	.13	.75	.16	.79	.25	.64	.20	.66	.23	.60	.06	.72	.14	.12	.06	.11	.14	.17	.22	.86
1974	.06	.84	.03	.89	.13	.78	.14	.95	.10	1.08	.07	.91	.06	.10	.08	.09	.09	-.17	.02	.94
1975	.13	.84	.17	.86	.14	.64	.20	.70	.28	.72	.05	.82	.20	.03	-.03	.22	.10	.00	.13	.80
1976	.10	.77	.13	.71	.10	.90	.19	.77	.28	.68	.10	.96	.24	.13	.16	.00	.04	.04	-.06	.76
1977	.15	.73	.15	.67	.18	.74	.19	.72	.35	.52	.09	.87	.28	.12	.18	.08	.09	.13	.04	.72
1978	.19	.69	.18	.59	.20	.55	.20	.57	.21	.41	.03	.60	.22	.12	.23	.25	.23	.39	.37	.78
1979	.19	.66	.18	.58	.26	.57	.25	.67	.32	.51	.10	.82	.10	.13	.25	.17	.08	.17	.08	.90
1980	.22	.70	.19	.59	.26	.60	.26	.63	.27	.57	.11	.86	.31	.08	.22	.15	.11	.15	.03	.69
1981	.20	.69	.17	.59	.25	.61	.24	.65	.24	.67	.11	.92	.31	.11	.24	.15	.11	.09	-.03	.69

* includes the graduates who have not pursued their studies in the public sector schools

NOTE: year t means school year t-1/t

SOURCE: Diagnostico/Previsoes, Educational Planning Bureau (GEP), Ministry of Education, Portugal, 1983

Table 4.7 School Enrolment*

Year	1960/61	1970/71	1977/78	1981/82
Educational enrolment	1 142 000	1 503 300	1 905 200	1 950 000
Indices of enrolment	100.00	131.64	166.83	170.75

*See 'Review of National Policies for Education - Portugal', OECD, Paris, 1984

of the increase observed in the first decade can be explained, as discussed in the previous section, by the intensification of female school attendance, so that nowadays the distribution of enrolment by sex is the same for males and females, not only in compulsory education but also in secondary education.

The basic Markov model described in the preceding chapter was summarised by equation (3.7) as follows:

$$\underline{n}^* = \underline{Np} + \underline{u}$$

When applying this model to the Portuguese educational subsystem on which this study concentrates, the number of states of the system is 7 and the number of observations per state is 11. States 1-7 correspond to grades 5-11 and state 8 gives the global drop-outs of the system. The dimensions of the model are described as follows:

$\underline{n}^*$ - is an (88 x 1) column vector whose first 77 elements are the number of students by observed year t in the seven different grades. The first eleven elements were replaced by the number of repeaters in the first grade of analysis as $\underline{n}^* = \underline{n} - \underline{n}_o$ is the vector of the number of students after deducting the new entrants to the system. The unavailability of the number of new-entrants to the school system in the different grades except the first one, imposes the condition that the $\underline{n}_o$ vector of new entrants has only eleven non-zero elements, the first ones, the remainder being equal to zero. The last eleven elements of the vector $\underline{n}^*$ reproduce the global drop-outs for the eleven observed school years.

$\underline{N}$ - is an (88 x 56) block diagonal matrix with eight matrices on the main diagonal. Each of these matrices is a (11 x 7) matrix of the lagged values of the number of students in different grades (observed values in year t-1).

$\underline{p}$ - is an (88 x 1) column vector of the transition probabilities in which some of the values are zero according to the structure presented in Table 4.5.

$\underline{u}$ - is an (88 x 1) vector of the residuals.

Equation (3.7) is a stacked form describing the relationship between the number of students in each grade at time t with the number of students in all grades at time t-1. These values have been stacked over time and over grades. Eliminating the zero values of the transition probabilities, the model described by equation (3.7) can be rewritten in a detailed form as:

$$
\underset{(88 \times 1)}{\begin{bmatrix} n_1^*(1) \\ \vdots \\ n_1^*(k) \\ n_2(1) \\ \vdots \\ n_2(k) \\ \vdots \\ n_8(1) \\ \vdots \\ n_8(k) \end{bmatrix}}
=
\underset{(88 \times 21)}{\begin{bmatrix} n_1(0) & & & & & & \\ \vdots & & & & & & \\ n_1(k-1) & & & & & & \\ & n_1(0) & n_2(0) & & & & \\ & \vdots & \vdots & & & & \\ & n_1(k-1) & n_2(k-1) & & & & \\ & & & \ddots & & & \\ & & & & n_1(0) & \cdots & n_7(0) \\ & & & & \vdots & & \vdots \\ & & & & n_1(k-1) & \cdots & n_7(k-1) \end{bmatrix}}
\underset{(21 \times 1)}{\begin{bmatrix} p_{11} \\ p_{12} \\ p_{22} \\ \vdots \\ p_{18} \\ p_{28} \\ \vdots \\ p_{78} \end{bmatrix}}
+
\underset{(88 \times 1)}{\begin{bmatrix} u_1(1) \\ \vdots \\ u_1(k) \\ u_2(1) \\ \vdots \\ u_2(k) \\ \vdots \\ u_8(1) \\ \vdots \\ u_8(k) \end{bmatrix}}
$$

Equations (3.5), however, describe the same relationship by grade. These equations, applied to the subsystem under analysis, can be written in a more disaggregated form. Replacing states 1-7 by grades 5-11 and state 8 by grade d, the eight equations of the model are as follows:

$$
\begin{array}{ll}
\text{EQ}_1 & \underline{n}_5^{*t} = \underline{n}_5^{t-1}\, p_{55} + \underline{u}_1 \\
\text{EQ2} & \underline{n}_6^{t} = \underline{n}_5^{t-1}\, p_{56} + \underline{n}_6^{t-1}\, p_{66} + \underline{u}_2 \\
\text{EQ3} & \underline{n}_7^{t} = \underline{n}_6^{t-1}\, p_{67} + \underline{n}_7^{t-1}\, p_{77} + \underline{u}_3 \\
\text{EQ4} & \underline{n}_8^{t} = \underline{n}_7^{t-1}\, p_{78} + \underline{n}_8^{t-1}\, p_{88} + \underline{u}_4 \\
\text{EQ5} & \underline{n}_9^{t} = \underline{n}_8^{t-1}\, p_{89} + \underline{n}_9^{t-1}\, p_{99} + \underline{u}_5 \\
\text{EQ6} & \underline{n}_{10}^{t} = \underline{n}_9^{t-1}\, p_{9,10} + \underline{n}_{10}^{t-1}\, p_{10,10} + \underline{u}_6 \\
\text{EQ10} & \underline{n}_{11}^{t} = \underline{n}_{10}^{t-1}\, p_{10,11} + \underline{n}_{11}^{t-1}\, p_{11,11} + \underline{u}_7 \\
\text{EQ8} & \underline{n}_d^{t} = \sum_{i=5}^{11} \underline{n}_i p_{id} + \underline{u}_8
\end{array}
\qquad (4.1)
$$

4.5. The Unrestricted OLS Estimator

The first step in the process of estimating the paramaters of the model was to obtain the unrestricted OLS estimates of the transition probabilities. The REGRESSION computational procedure, included in the SPSS package, version 7, available at the University of Manchester Regional Computer Centre, was used (see program SPSSREG included in Appendix B). The multiple regression procedure gives the statistics necessary to evaluate the results and is adjustable to the purposes of this study as it has an option that enables the forcing of the regression through the origin. The estimated values are presented in Table 4.8 which shows that the non-negativity condition on the set of the transition probabilities and the row-sum condition are both violated.

When comparing the estimates of the transition probabilities with the corresponding observed point estimates, some of them ($\hat{p}_{66}$, $\hat{p}_{89}$, $\hat{p}_{10,10}$) present very high values and others ($\hat{p}_{56}$, $\hat{p}_{99}$, $\hat{p}_{9,10}$) present very low values. Also, only a few of the estimated probabilities show low standard errors, thus causing some of the 95% confidence intervals to be very large. However, even though large, some of these confidence intervals do not contain most of the observed point estimates.

It is important to note at this stage that grades 6 and 9 are terminal grades of levels of education (preparatory and secondary unified general course), affected by global evaluations at the end of the academic year. Also, grades 7 and 10, especially grade 10, absorb a significant number of new entrants coming from the private schools every year. At the same time, a high number of drop-outs

Table 4.8

The Unrestricted OLS Estimates of the Transition Probabilities

Transition Probability	Estimated Value	Standard error	t-value*	95% Confidence Interval	
				L.B.	U.B.
p_{55}	0.16	0.0146	10.938	0.128	0.193
p_{56}	0.59	0.1792	3.295	0.185	0.996
p_{66}	0.33	0.2215	1.494	-1.170	0.832
p_{67}	0.57	0.3151	1.798	-0.147	1.278
p_{77}	0.34	0.3949	0.855	-0.556	1.231
p_{78}	0.72	0.1856	3.901	0.304	1.144
p_{88}	0.15	0.2282	0.650	-0.368	0.665
p_{89}	1.01	0.1214	8.339	0.738	1.287
p_{99}	-0.12	0.1386	0.832	-0.428	0.198
$p_{9,10}$	0.06	0.1897	0.335	-0.366	0.492
$p_{10,10}$	0.95	0.3253	2.917	0.213	1.685
$p_{10,11}$	0.81	0.2142	3.775	0.324	1.293
$p_{11,11}$	0.23	0.2280	1.127	-0.259	0.773
p_{5d}	2.33	0.8305	2.806	0.025	4.634
p_{6d}	-1.43	1.1326	1.263	-4.575	1.715
p_{7d}	0.36	0.6065	0.600	-1.320	2.048
p_{8d}	2.77	0.6297	4.400	1.022	4.519
p_{9d}	2.10	1.4855	1.417	-2.019	6.230
$p_{10,d}$	-0.89	0.7005	1.266	-2.832	1.058
$p_{11,d}$	-3.37	1.6184	2.084	-7.856	1.121

* $t^{9}_{0.025} = 2.262$; $t^{8}_{0.025} = 2.306$; $t^{3}_{0.025} = 3.108$

take place at the end of these levels, not properly shown by the drop-out probabilities, as these compensate the drop-outs with the new entrants to the grade. A look at the estimates of the transition probabilities shows that those most at variance with the point estimates are the estimates concerning the link between two levels of education. A special reference to the drop-out estimates should also be made. When the non-negativity condition is not taken into account, only one drop-out probability estimate falls into the acceptable range 0-1.

Knowing that the OLS estimators have desirable properties (unbiasedness with the smallest variance) is only cold comfort if their variances are such that the resulting estimates are highly unreliable. That is, knowing that the estimator has the smallest possible variance (among all the unbiased estimators) is not very helpful if, at the same time, this variance happens to be very large. The basic assumptions of the regression model require that none of the exogenous variables be perfectly correlated with any other exogenous variable, or with any linear combination of the exogenous variables. A case of a high degree of multicollinearity arises whenever one exogenous variable is highly correlated with another exogenous variable, or with a linear combination of other exogenous variables.

The characteristic of the basic Markov model presupposes a high degree of multicollinearity between the independent variables explained by the high values observed for the standard errors. When multicollinearity exists, there is not too much to do about it except to make changes either in the method of estimation, if it is worth making such changes, or to make rearrangements in the number of explanatory variables included. The exclusion of any explanatory variable from this model is unrealistic.

The F-distribution for each of the equations of the model, as presented in Table 4.9, determines whether or not all the partial regression coefficients are equal to zero. High F-statistic values with 5% significance were found, which suggests the acceptance of the regression as a whole, indicating a strong relationship between school enrolment in a certain grade at time t and the school enrolment in the same grade and in the preceding one at time t-1.

The reliability of the partial regression estimates for each equation can be examined through the t-values. If the calculated t-value is greater, in absolute value, than the critical value at the 5% level of significance, the hypothesis of no relationship between the corresponding explanatory variable and the dependent variable is to be rejected. The t-values presented in Table 4.8 show that for all equations with two estimated coefficients, only one of them is greater than the critical value. Also, for EQ8 which estimates the drop-out probabilities, only $\hat{p}_{8d}$ is significant at the 5% level. However, the estimated value is out of the range 0-1 and the corresponding 95% confidence interval does not contain any of the observed point estimates. Therefore, it is reasonable to have some doubts about the reliability of the unrestricted OLS estimates of the transition probabilities for the basic Markov model.

The degree to which changes in the set of the independent variables generate changes in the dependent variable is measured by the coefficient of multiple determination or R^2. When the model has no constant term, that is, when the regression is forced through the

Table 4.9

Coefficients of Determination, F-Statistics and Durbin-Watson Statistics for the Unrestricted OLS Estimator

Equation	R^2	F	D-W Statistic	Coefficient of variability
EQ1	0.4683	119.64	0.729	0.32
EQ2	0.9233	2104.32	1.321	0.05
EQ3	0.7143	460.49	0.828	0.11
EQ4	0.7896	314.80	1.828	0.14
EQ5	0.9425	679.18	1.919	0.10
EQ6	0.6816	137.51	2.315	0.21
EQ7	0.9050	250.27	1.392	0.16
EQ8	0.9552	59.29	2.363	0.18

NOTE 1: $F_{0.05}(1,9)=5.12$; $F_{0.05}(2,8)= 4.46$; $F_{0.05}(7,3)= 8.89$

NOTE 2 : $DW_{0.05}(11,1)=(0.768,2.511)$; $DW_{0.05}(11,2)=(0.610,2.231)$;
$DW_{0.05}(11,7)=(0.058,0.567)$ [see FAREBROTHER, 1980]

origin, an alternative value of R^2 proposed by THEIL [1971] should be calculated. The formula is $R^2 = (SSTa - SSEu)/SSTa$ where SSTa is the total sum of squares adjusted for the mean and SSEu is the error sum of squares unadjusted for the mean. R^2 provides information about the explanatory power of the model. The subprogram REGRESSION used gives this alternative value of R^2. However, in dealing with time series data, very high R^2s are not unusual due to common trends. AMES and REITER [1961] found, for example, that on average, the R^2 of a relationship between a randomly chosen variable and its own value lagged one period is about 0.7 and that an R^2 greater than 0.5 could be obtained by selecting an economic time-series and regressing it against two to six other randomly selected economic time-series. In this case, the "good" values for the coefficients of determination presented in Table 4.9 are not unexpected.

One of the assumptions underlying the regression analysis of time-series data is that the error terms of the different observations of the variables are not related. But when this is not true the problem of autocorrelation occurs. The major problem with autocorrelation is that it may cause the researcher to accept a partial regression coefficient as being significantly different from zero when it is not, and secondly it may cause acceptance of the null hypothesis that the partial regression coefficient is zero when it is different from zero. This arises because the t-test generates a t-statistic which is greater than the critical value in the first case and smaller in the second case. The REGRESSION program used provides the Durbin-Watson statistics which indicates the existence or not of the autocorrelation. The values presented in Table 4.9 reveal that only two equations show clearly the existence of no autocorrelation.

Finally, the standard error of the estimate provides information about the predictive power of the model. This statistic is directly related to the stochastic term u in the equation. The program used, in addition to generating information about the estimate of the error term, also provides the standard error of the estimate. It measures the spread of the data around the estimated regression line. However, the size on its own is of no importance; it must be examined with reference to the size of the mean of the dependent variable. The ratio of the standard error of the estimate to the mean of the dependent variable is called the coefficient of variability and its values are presented in the last column of Table 4.9. The smaller is this value, the greater is the predictive power of the model.

The analysis so far is based on the simplifying assumption that heteroscedasticity is not present and that the error term is normally distributed with zero mean. The error region is then bounded by straight lines; this is not true. Rather than being parallel straight lines, the boundaries of the error region are curved. The size of the error grows as one move away from the mean of the variables [see PINDYCK and RUBINFELD (1981),p.211]. It is also why predictions of the dependent variable made from independent variable values far from the mean have a much greater probability of being incorrect. This is the case of our model as the number of students by grade present standard deviations slightly bigger than the corresponding means, which causes high values for the coefficients of variability. However, it is not the intention of this study to use a Markov model for predictive purposes. As already discussed, the aim of the study is to get a better understanding of the behaviour of the regression coefficients (the transition probabilities).

4.6. The Restricted OLS Estimator

As seen in the previous section, when applying the unrestricted OLS estimator to the set of equations of the model, the non-negativity and row-sum conditions are both violated. The inclusion of these conditions into the model turns the problem of estimating the transition probabilities, as shown in Chapter 3, into the following constrained quadratic programming problem

$$\min_{\underline{\hat{p}}} \; (\underline{n}^* - \underline{N\hat{p}})' \; (\underline{n}^* - \underline{N\hat{p}})$$

$$\text{s.t.} \quad \underline{G}\,\underline{\hat{p}} = \underline{n}_{s+1}$$

$$\underline{\hat{p}} > \underline{0}$$

Thus, the first step on which this section concentrated has been the building of the objective function. For ease of computation, the transition probabilities listed in Table 4.8 as p_{55} to $p_{11,d}$ were renamed as V1 to V20 in the same sequential order. Auxilliary FORTRAN programs JOB(M) and JOB(MMM) were written, and used to construct the objective function. The minimisation problem (see program MPOS1, Appendix B) was solved using the SYMQUAD computational procedure available in the integrated system of computer programs MPOS (Multipurpose Optimization System). Among the four quadratic algorithms available in this system of programs (WOLFE, BEALE, LEMKE and SYMQUAD), SYMQUAD has been chosen, as experience has shown that this algorithm appears to be superior to the others, as well as faster.

As MPOS package programs do not provide the statistics for the optimum values of the variables, FORTRAN program RESIDD was written in order to obtain the statistics for the repetition and promotion probabilities. These values are presented in Table 4.10. The statistics for the drop-out probabilities have not been computed as the attention of this study is focused on the behaviour of the repetition and promotion probabilities. Knowing that the non-negativity and row-sum conditions are imposed on the model, the drop-out probability estimates can easily be obtained by using the row-sum condition. However, it seems relevant to note that, after applying the row sum condition, three drop-out probability estimates ($\hat{p}_{5d}$, $\hat{p}_{7d}$, $\hat{p}_{10d}$) are equal to zero; also, two of the four non-null values for these estimates ($\hat{p}_{8d}$, $\hat{p}_{9d}$) have the corresponding repetition probability estimates equal to zero. This is unreal as the observed point estimates of the different probabilities (see Table 4.6) for the time period of analysis are different from zero.

Globally, the restricted OLS estimator appears to generate more accurate estimates of the transition probabilities than the unrestricted OLS estimator; the standard errors of the estimates are smaller, and consequently the 95% confidence intervals for the restricted OLS estimates are less wide than those for the unrestricted OLS estimates. However, the estimates for the transition probabilities associated with grades beyond grade 8 are very different from the observed point estimates. In particular, the observed point estimates for the transition probabilities $p_{10,10}$ and $p_{10,11}$ do not fall within the bounds of the corresponding 95% confidence intervals.

While testing the estimates of the transition probabilities using the t-values, most of the equations show that only one of the independent variables has a high statistical significance (see Table 4.10). The same results were obtained for the unrestricted OLS

Table 4.10

The Restricted OLS Estimates of the Transition Probabilities

Transition Probability	Estimated Value	Standard error	t-value*	95% Confidence Interval	
				L.B.	U.B.
p_{55}	0.16	0.0863	1.896	-0.032	0.358
p_{56}	0.84	0.0873	9.622	0.635	1.104
p_{66}	0.03	0.0836	0.359	-0.167	0.218
p_{67}	0.73	0.1515	4.818	0.383	1.108
p_{77}	0.14	0.1363	1.053	-0.171	0.458
p_{78}	0.86	0.1421	6.052	0.529	1.118
p_{88}	0.00	0.1506	0.000	-0.347	0.347
p_{89}	0.93	0.0785	11.847	0.746	1.111
p_{99}	0.00	0.0913	0.000	-0.211	0.211
$p_{9,10}$	0.25	0.1873	1.335	-0.184	0.679
$p_{10,10}$	0.58	0.1016	5.729	0.347	0.817
$p_{10,11}$	0.42	0.1344	3.109	0.108	0.728
$p_{11,11}$	0.59	0.1630	3.595	0.210	0.962

* $t^{9}_{0.025}$ =2.262; $t^{8}_{0.025}$ = 2.306

Table 4.11

The F-Statistic for the Restricted OLS Estimator

Equation	F
EQ1	54.62
EQ2	608.47
EQ3	756.47
EQ4	1118.87
EQ5	3501.34
EQ6	1683.46
EQ7	2954.97

*$F_{0.05}(1,9)$=5.12; $F_{0.05}(2,8)$= 4.46

estimator and are not unexpected due to the high degree of multicollinearity between the independent variables.

In almost all of the regressions, higher F-statistic values, significant at the 5% level, were found when compared to the F-statistic values obtained for the unrestricted OLS estimator.

4.7. The Unrestricted GLS Estimator

In order to obtain efficient estimators of the transition probabilities, the generalised least squares (GLS) estimation procedure was used. This method is equivalent to applying the OLS estimation process to transformed data. The assumption of efficiency established for the OLS estimator states that the error disturbances are normally distributed with constant variances over observations. This is not the case for the model for most of the equations; therefore, the error disturbances are likely to be heteroscedastic, that is the variances of the error disturbances are not constant over observations. When heteroscedasticity is present, OLS estimation places more weight on the observations which have larger error variances than on those with small variances. The implicit weighting of OLS occurs because the sum of squared residuals associated with large variance error terms is likely to be substantially greater than the sum of squared residuals associated with small variance errors. Because of this implicit weighting, the study is not only dealing with a high degree of multicollinearity, but also with heteroscedastic error disturbances. The variances of the estimated transition probabilities are not the minimum variances. Further, they are biased estimators of the true variances of the estimated probabilities.

In order to correct for heteroscedasticity, the assumption of a constant variance for the disturbance term is now replaced by the assumption that the variance - covariance matrix of the disturbances is known and has the form

$$\underline{\Sigma} = E(\underline{u}\,\underline{u}') = \begin{bmatrix} E(\underline{u}_1\,\underline{u}_1') & \cdots & E(\underline{u}_7\,\underline{u}_7') \\ \cdots & \cdots & \cdots \\ E(\underline{u}_7\,\underline{u}_1') & \cdots & E(\underline{u}_7\,\underline{u}_7') \end{bmatrix}$$

Each term of the principal diagonal of the $\underline{\Sigma}$ matrix is the (11 x 11) variance-covariance matrix of the residuals for each of the equations of the model. The off-diagonal terms of $\underline{\Sigma}$ represent the (11 x 11) matrices whose elements are the contemporaneous and lagged covariances between disturbances from a pair of equations. By assumption, $E(\underline{u}_i\,\underline{u}_j') = \sigma_{ij}\underline{I}$ (i=1, ... ,7). Therefore the $\underline{\Sigma}$ matrix can be written as follows

$$\underline{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{17} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{27} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{71} & \sigma_{71} & \cdots & \sigma_{77} \end{bmatrix} \otimes \underline{I}$$

where $\otimes$ denotes the Kronecker multiplication of matrices and $\underline{I}$ is an identity matrix of order 11. $\underline{\Sigma}$ is then the symmetric matrix of size (77 x 77).

The principal difficulty in applying the generalised least squares method is that $\underline{\Sigma}$ is unknown. To overcome this point, ZELLNER [1962] proposes that ordinary least squares be applied to each equation and the computed residuals used to estimate the elements of $\underline{\Sigma}$.[4] FORTRAN program SIGMA was written to compute the $\underline{\Sigma}$

matrix, and its inverse $\underline{\Sigma}^{-1}$, using the residuals obtained by applying the OLS estimation procedure and estimating σ_{ii} and σ_{ij} using the following expressions

$$s_{ii} = \underline{u}_i' \, \underline{u}_i / n-k_i$$

$$s_{ij} = \underline{u}_i' \, \underline{u}_j / (n-k_i)^{1/2} \, (n-k_j)^{1/2}$$

where $n = 11$, $k_i = 1$ and $k_j = 2$, $j = 2, \ldots, 7$. These estimates were substituted in $\underline{\Sigma}$ to obtain the estimated matrix $\underline{\Sigma}^{-1}$. In this section, the last equation of the model was omitted, the estimates of the drop-out probabilities being obtained a posteriori using the row-sum condition.

FORTRAN program UNGLS computed the unrestricted GLS transition probability estimates using expression (3.16)

$$\underline{\hat{p}} = (\underline{N}' \, \underline{\Sigma}^{-1} \underline{N})^{-1} \, (\underline{N}' \, \underline{\Sigma}^{-1} \underline{n}^*)$$

where $\underline{N}' \, \underline{\Sigma}^{-1} \underline{N}$ is a symmetric matrix of size (13 x 13) with the following form

$$\underline{N}' \, \underline{\Sigma}^{-1} \underline{N} = \begin{bmatrix} s^{11}\underline{N}_1'\underline{N}_1 & s^{12}\underline{N}_1'\underline{N}_2 & \cdots & s^{17}\underline{N}_1'\underline{N}_7 \\ s^{21}\underline{N}_2'\underline{N}_1 & s^{22}\underline{N}_2'\underline{N}_2 & \cdots & s^{27}\underline{N}_2'\underline{N}_7 \\ \cdots & \cdots & \cdots & \cdots \\ s^{71}\underline{N}_7'\underline{N}_1 & s^{72}\underline{N}_7'\underline{N}_2 & \cdots & s^{77}\underline{N}_7'\underline{N}_7 \end{bmatrix}$$

where $\underline{N}_1$ is a vector of size (11 x 1), $\underline{N}_i$ $(i=2, \ldots, 7)$ are matrices of size (11 x 2) and s^{ij} are the elements of matrix $\underline{\Sigma}^{-1}$. Vector $(\underline{N}' \, \underline{\Sigma}^{-1} \underline{n}^*)$ is the (13 x 1) with the form

$$\underline{N}'\,\underline{\Sigma}^{-1}\underline{n}^{*} = \begin{bmatrix} \sum_{i=1}^{7} s^{1i}\ \underline{N}_1'\ \underline{n}_i^{*} \\ \sum_{i=1}^{7} s^{2i}\ \underline{N}_2'\ \underline{n}_i^{*} \\ \vdots \\ \sum_{i=1}^{7} s^{7i}\ \underline{N}_7'\ \underline{n}_i^{*} \end{bmatrix}$$

The estimates of the transition probabilities, the standard errors, t-values and 95% confidence intervals are presented in Table 4.12. As in the case of the restricted OLS model, the statistics of the transition probability estimates were computed applying program RESIDD to the new estimates obtained.

The unrestricted GLS estimator does not violate the non-negativity condition for the repetition and promotion probabilities, but the row-sum condition is not satisfied for these equations (EQ2, EQ4, EQ7). In statistical terms, the GLS regression estimator is recomended over the OLS estimator. However, the GLS regression estimator does not give any closer estimates, when comparing these values with the time series of the observed point estimates of the transition probabilities presented in Table 4.6.

In general, the unrestricted GLS estimates of the transition probabilities show smaller standard errors than the standard errors obtained when performing the unrestricted OLS estimator. Thus, in all cases but two, the observed point estimates of the transition probabilities fall within the bounds of the 95% confidence interval for the unrestricted GLS estimator; $\hat{p}_{55}$ and $\hat{p}_{56}$ are very far from the point estimates, so that not even the 99% confidence interval ($0.179 < \hat{p}_{56} < 0.672$ and $0.284 < \hat{p}_{66} < 0.757$) contains the observed point estimates. The restricted OLS estimates for these probabilities are

Table 4.12

The Unrestricted GLS Estimates of the Transition Probabilities

Transition Probability	Estimated Value	Standard error	t-value*	95% Confidence Interval	
				L.B.	U.B.
p_{55}	0.19	0.1012	1.864	-0.004	0.417
p_{56}	0.43	0.0735	6.169	0.256	0.595
p_{66}	0.52	0.0704	7.392	0.358	0.683
p_{67}	0.81	0.1538	5.257	0.454	1.163
p_{77}	0.01	0.1388	0.098	-0.306	0.334
p_{78}	0.43	0.1976	2.176	-0.026	0.885
p_{88}	0.39	0.2095	1.875	-0.090	0.876
p_{89}	0.78	0.0905	0.086	0.572	0.990
p_{99}	0.11	0.1052	9.772	-0.136	0.350
$p_{9,10}$	0.55	0.2346	2.343	0.009	1.091
$p_{10,10}$	0.06	0.1273	0.442	-0.237	0.450
$p_{10,11}$	0.49	0.6538	0.743	-1.021	1.994
$p_{11,11}$	1.50	0.7932	1.891	-0.258	3.332

$*t^{9}_{0.025} = 2.262$; $t^{8}_{0.025} = 2.306$

Table 4.13

The F-Statistic for the Unrestricted GLS Estimator

Equation	F
EQ1	45.76
EQ2	644.60
EQ3	720.83
EQ4	478.29
EQ5	2443.22
EQ6	953.50
EQ7	233.28

$*F_{0.05}\ (1,9)=5.12$; $F_{0.05}\ (2,8)= 4.46$

more accurate and even the 90% confidence interval ($0.674 < \hat{p}_{56} < 0.999$ and $-0.130 < \hat{p}_{66} < 0.181$) contain practically all of the corresponding observed point estimates. A comparison between Table 4.12 and Table 4.10 shows that, with the exception of $\hat{p}_{56}$ and $\hat{p}_{66}$, the remaining standard errors for the restricted OLS estimator are smaller than those obtained for the corresponding unrestricted GLS estimator, indicating this is therefore more efficient.

An attempt to compute the restricted GLS estimators was made. As for the case of the OLS estimation procedure, when the non-negativity and row-sum conditions are forced into the model, a quadratic programming problem occurs. The problem to be solved is, as previously discussed in Chapter 3 and Appendix A, given by expressions (A.11) and (A.12),

$$\min_{\underline{\hat{p}}} (\underline{n}^* - \underline{N}\underline{\hat{p}})' \, \underline{\Sigma}^{-1} (\underline{n}^* - \underline{N}\underline{\hat{p}})$$

$$\text{s.t} \quad \underline{R}\,\underline{\hat{p}} < \underline{n}_{s+1}$$

$$\underline{\hat{p}} > \underline{0}$$

FORTRAN program SIGMA was applied to the residuals of the restricted OLS estimation procedure in order to obtain the variance-covariance matrix $\underline{\Sigma}$ of the disturbance terms and its inverse. The objective function $(\underline{n}^* - \underline{N}\,\underline{\hat{p}})' \, \underline{\Sigma}^{-1} (\underline{n}^* - \underline{N}\,\underline{\hat{p}})$ was then constructed and the quadratic programming algorithm SIMQUAD from the MPOS package was applied (see FORTRAN program SIGMA2 and program MPOS4 in Appendix B). Unfortunately no minimum was obtained for the objective function, as when the constraints are imposed the algorithm sets the first variable to the upper bound (1.0), the second variable to the lower bound (0.0) and returns the remaining variables as non-basic.[5] This shows that the inclusion of the constraint (row-sum condition) has led to a non-positive definite (or non-positive semidefinite) quadratic form for the objective function.

Before proceeding with this study, a comparative analysis between the estimates obtained by the three estimation procedures should be made. Table 4.14 summarises the results obtained by comparing these estimates with the observed point estimates of the transition probabilities. All the estimation methods are non-efficient and unreliable. However, the restricted OLS estimator seems to be the closer to the observed point estimates of the transition probabilities. Even so, these estimates are far from being acceptable and reliable transition probability estimates for the Portuguese educational system. The observed point estimates present strong changes over the time period of analysis. It seems that the assumption of constant transition probabilities over time is quite limited and unreal.

4.8. The Estimate of the Returnees and the OLS Estimator Applied to the Smoothed Data

The figures given in Table 4.4 show that during the period of analysis covered by this study there were strong fluctuations in the observed point estimates of the transition probabilities. In the previous sections the use of the basic Markov model showed a non-stability of the transition probabilities and consequently led to non-efficient estimators for these probabilities. Before concluding that the model cannot be applied to the Portuguese educational system, a thorough examination of the fluctuations was conducted. Obviously in practice most transition rates tend to change. However, when these changes are substantial, they have to be taken into

Table 4.14 The Observed Point Estimates and the OLS and GLS Estimates For the Transition Probabilities - Whole Country

Trans Probab.	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,1}$	P_{5d}	P_{6d}	P_{7d}	P_{8d}	P_{9d}	$P_{10,d}$	$P_{11,d}$
VAR CODE / YEAR	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20
1971	.16	.67	.17	.63	.23	.67	.26	.58	.21	.58	.04	.77	.11	.17	.19	.10	.16	.21	.23	.89
1972	.15	.73	.16	.66	.22	.68	.22	.64	.23	.64	.06	.77	.12	.12	.18	.10	.14	.12	.20	.88
1973	.13	.75	.16	.77	.25	.64	.20	.66	.23	.60	.06	.72	.14	.12	.06	.11	.14	.17	.22	.86
1974	.06	.84	.03	.89	.13	.78	.14	.95	.10	1.08	.07	.91	.06	.10	.08	.09	.09	-.17	.02	.94
1975	.13	.84	.17	.86	.14	.64	.20	.70	.28	.72	.05	.82	.20	.03	-.02	.22	.10	.00	.13	.80
1976	.10	.77	.13	.71	.10	.90	.19	.77	.28	.68	.10	.96	.24	.13	.16	.00	.04	.04	-.06	.76
1977	.15	.73	.15	.67	.18	.74	.19	.72	.35	.52	.09	.87	.28	.12	.18	.08	.09	.13	.04	.72
1978	.19	.69	.18	.59	.20	.55	.20	.57	.21	.41	.03	.60	.22	.12	.23	.25	.23	.39	.37	.78
1979	.19	.66	.18	.58	.26	.57	.25	.67	.32	.51	.10	.82	.10	.13	.25	.17	.08	.17	.08	.90
1980	.22	.70	.19	.59	.26	.60	.26	.63	.27	.57	.11	.86	.31	.08	.22	.15	.11	.15	.03	.69
1981	.20	.69	.17	.59	.25	.61	.24	.65	.24	.67	.11	.92	.31	.11	.24	.15	.11	.09	-.03	.69
ESTIMATES																				
Unr.OLS*	.16	.59	.33	.57	.34	.72	.15	1.01	-.12	.06	.95	.81	.23	.25	.10	-.06	-.16	1.06	-.76	.77
Rest.OLS	.16	.84	.03	.73	.14	.86	.00	.93	.00	.25	.58	.42	.59	.00	.24	.00	.07	.75	.00	.49
Unr.GLS*	.19	.43	.52	.81	.01	.43	.39	.78	.11	.55	.06	.49	1.50	.38	-.33	.56	-.17	.34	.45	-.50

*The drop-out probabilities were calculated by difference, assuming the row-sum condition is satisfied; they are not the OLS estimates account when preparing enrolment projections or analysing the behaviour of the school enrolment. These changes may be the result of policy measures, of the introduction of new laws concerning compulsory school attendance, or of changes in the structure of the school system. The twelve year period covered by the study, as well as the delimitation of the two educational levels (preparatory and secondary) were chosen in order to avoid an overlap with any measures which can bring about these changes. As already mentioned, during this period the Portuguese educational system, particularly the public sector school system, was faced with the problem of accommodating the students transfered from Angola and Mozambique. So far in this section it has been assumed that the primary reason for the observed disturbances in the point estimates of the transition probabilities was the increase in the school enrolment after the revolution, due to the returnees from Africa.

As discussed in section 4.4, the allocation of the returnee students into the school system took place without any control by the Ministry of Education. Thus, no data concerning quantitative or qualitative aspects of the returnee enrolment are available. It is also unknown during which school years the allocation procedure was carried out. This unavailability of data makes difficult any attempt to get reasonable estimates of the number of the returnee students. Nevertheless, to overcome the problem the following hypotheses were established:

(i) The allocation of the returnee students into the public school system took place in the school years 1974/75 and 1975/76.

(ii) Before the revolution, a stable behaviour was presented by the educational system, manifested by constant probabilities whose values are the mean values of the corresponding observed point estimates in 1971, 1972 and 1973.

An iterative process was used to: (1) generate the two matrices of returnee enrolment, one referring to the entrants to the school system in the school year 1974/75 and the other referring to those who entered in 1975/76; (2) Compute the smoothed matrix of enrolment; (3) give the OLS estimators of the transition probabilities.

Iterative Process

Step 1 - Using the constant transition probabilities referred to in (ii), the number of returnee students allocated by the public school system in 1974/75 and 1975/76 are estimated using the following expressions:

$$r_i^{74} = n_i^{74} - (n_{i-1}^{73}\, p_{i-1,i} + n_i^{74}\, p_{ii})$$

$$r_i^{75} = n_i^{75} - r_{i,e}^{75} - (n_{i-1}^{74} - r_{i-1}^{74})p_{i-1,i} - (n_i^{74} - r_i^{74})p_{ii}$$

$$i = 5, \ldots, 11$$

where

r_i^t is the number of returnee students entering the public school system to grade i, in the school year t/t+1.

$r_{i,e}^{75}$ is the number of returnee students who, having entered the public school system in 1974/75, are in grade i in 1975/76.

n_i^t are the observed values of school enrolment presented in Table 4.4.

$p_{i-1,i}$ is the constant transition probability from grade i-1 to grade i.

Step 2 - The values estimated in Step 1 are the first rows of the matrices of returnee enrolment. Denoting by $\underline{R}$ the matrix corresponding to the entrants in 1974/75 and denoting by $\underline{Q}$ the matrix corresponding to the entrants in 1975/76, the remaining rows of these two matrices are generated using the same constant transition probabilities referred to in (ii).

Step 3 - Sum $\underline{R}$ + $\underline{Q}$ to obtain the matrix $\underline{S}$ of the overall returnee enrolment.

Step 4 - Subtract $\underline{S}$ from $\underline{N}$ (the matrix of the observed values of enrolment presented in Table 3.4) in order to obtain the smoothed matrix of enrolment $\underline{A}$.

Step 5 - OLS estimators are applied to matrix $\underline{S}$ and matrix $\underline{A}$ in order to obtain estimates for the transition probabilities.

Step 6 - Go to Step 2 using the new estimates of the transition probabilities, obtained in Step 5, for the matrix $\underline{S}$ of returnee enrolment.

Program RETUR1 was written and used to perform Step 1 to Step 4 of the iterative process and provide a smoother matrix of the school enrolment. The estimated matrices of the returnee students allocated in 1974/75 and 1975/76 are presented in Appendix B, Tables B.1 - B.3 and the inferred estimated smoothed matrix $\underline{A}$ of enrolment is presented in Table B.4. Table B.5 shows the constant transition probabilities used and Table B.6 presents the estimated repeaters in grade 5 and the estimated drop-outs for the smoothed matrix of enrolment. The OLS estimation procedure was then performed to produce the estimates for the transition probabilities corresponding to matrices $\underline{S}$ and $\underline{A}$. The unrestricted and the restricted OLS estimates for the smoothed matrix $\underline{A}$ and the respective standard errors and 95% confidence intervals are presented in Table 4.15 and Table 4.17. Using the restricted OLS estimates of the transition probabilities for the matrix $\underline{S}$ of returnee enrolment, new matrices of returnee enrolment $\underline{R1}$ and $\underline{Q1}$ were computed. From these, new matrices for the returnee students (matrix $\underline{S1}$) and for the adjusted enrolment (matrix $\underline{A1}$) were obtained (see Appendix B, Tables B.7-B.10). The restricted OLS estimator was applied to the new smoothed data matrix $\underline{A1}$ and the results are shown in Table 4.19. A comparison between the different estimates obtained shows that although the unrestricted OLS estimation procedure does not satisfy the row-sum and the non-negativity conditions, it is the one which presents estimates most similar to the assumed constant repetition and transition probability values. However, the 95% confidence intervals for $\hat{p}_{9,10}$ and $\hat{p}_{10,10}$ do not contain the observed mean values. Table 4.21 summarises all the results obtained in this chapter.

The restricted OLS estimates of the parameters of the smoothed matrices of enrolment $\underline{A}$ and $\underline{A1}$ do not differ greatly from each other,

Table 4.15

The Unrestricted OLS Estimator for the Smoothed Matrix of Enrolments (Matrix A)

Transition Probability	Estimated Value	Standard error	t-value*	95% Confidence Interval	
				L.B.	U.B.
p_{55}	0.16	0.0150	10.850	0.129	0.196
p_{56}	0.72	0.1257	5.724	0.435	1.004
p_{66}	0.16	0.1590	0.987	-0.203	0.517
p_{67}	0.62	0.2481	2.489	0.056	1.179
p_{77}	0.25	0.3240	0.774	-0.482	0.984
p_{78}	0.65	0.1654	3.946	0.279	1.027
p_{88}	0.23	0.2069	1.115	-0.237	0.698
p_{89}	0.88	0.1270	6.939	0.594	1.169
p_{99}	0.02	0.1482	0.108	-0.319	0.351
$p_{9,10}$	-0.01	0.1367	0.058	-0.317	0.301
$p_{10,10}$	1.09	0.2528	4.297	0.514	1.658
$p_{10,11}$	0.72	0.3639	1.976	-0.104	1.542
$p_{11,11}$	0.37	0.3830	0.963	-0.498	1.235

Table 4.16

The F-Statistic for the Unrestricted OLS Estimator Applied to Data Matrix A

Equation	F
EQ1	117.72
EQ2	5030.48
EQ3	700.00
EQ4	262.98
EQ5	380.74
EQ6	270.75
EQ7	135.18

Table 4.17

The Restricted OLS Estimator for the Smoothed Matrix of Enrolments (Matrix $\underline{A}$)

Transition Probability	Estimated Value	Standard error	t-value*	95% Confidence Interval	
				L.B.	U.B.
p_{55}	0.18	0.0902	2.018	-0.250	0.358
p_{56}	0.59	0.0531	11.100	0.467	0.709
p_{66}	0.33	0.0464	7.107	0.234	0.442
p_{67}	0.66	0.1408	4.687	0.337	0.987
p_{77}	0.22	0.1125	1.956	-0.043	0.474
p_{78}	0.79	0.1758	4.494	0.378	1.119
p_{88}	0.09	0.1902	0.473	-0.035	0.513
p_{89}	0.91	0.1020	8.922	0.673	1.143
p_{99}	0.00	0.1251	0.000	-0.289	0.289
$p_{9,10}$	0.41	0.2081	1.971	-0.075	0.885
$p_{10,10}$	0.24	0.0989	2.427	0.015	0.473
$p_{10,11}$	0.76	0.1556	4.884	0.396	1.112
$p_{11,11}$	0.27	0.2029	1.321	-0.203	0.773

Table 4.18

The F-Statistic for the Restricted OLS Estimator Applied to Data Matrix $\underline{A}$

Equation	F
EQ1	52.28
EQ2	1738.99
EQ3	1258.04
EQ4	992.63
EQ5	2474.02
EQ6	1823.89
EQ7	3306.04

Table 4.19

The Restricted OLS Estimator for the Smoothed Matrix of Enrolments (Matrix A1)

Transition Probability	Estimated Value	Standard error	t-value*	95% Confidence Interval	
				L.B.	U.B.
p_{55}	0.18	0.0951	1.893	-0.032	0.400
p_{56}	0.58	0.0478	12.134	0.469	0.688
p_{66}	0.35	0.0433	8.077	0.250	0.448
p_{67}	0.65	0.1402	4.643	0.329	0.973
p_{77}	0.21	0.1069	1.966	-0.034	0.461
p_{78}	0.79	0.1862	4.243	0.354	1.121
p_{88}	0.09	0.1911	0.471	-0.364	0.546
p_{89}	0.91	0.0989	9.192	0.680	1.137
p_{99}	0.00	0.1242	0.000	-0.286	0.286
$p_{9,10}$	0.54	0.2257	2.323	0.023	1.066
$p_{10,10}$	0.00	0.1112	0.000	-0.256	0.256
$p_{10,11}$	1.00	0.1683	5.944	0.613	1.386
$p_{11,11}$	0.00	0.2121	0.000	-0.488	0.488

Table 4.20

The F-Statistic for the Restricted OLS Estimator Applied to Data Matrix A1

Equation	F
EQ1	50.63
EQ2	2030.16
EQ3	1312.05
EQ4	958.69
EQ5	2715.90
EQ6	1432.03
EQ7	3172.83

Table 4.21

The Observed Point Estimates for the Transition Probabilities and the Corresponding OLS and GLS Estimators

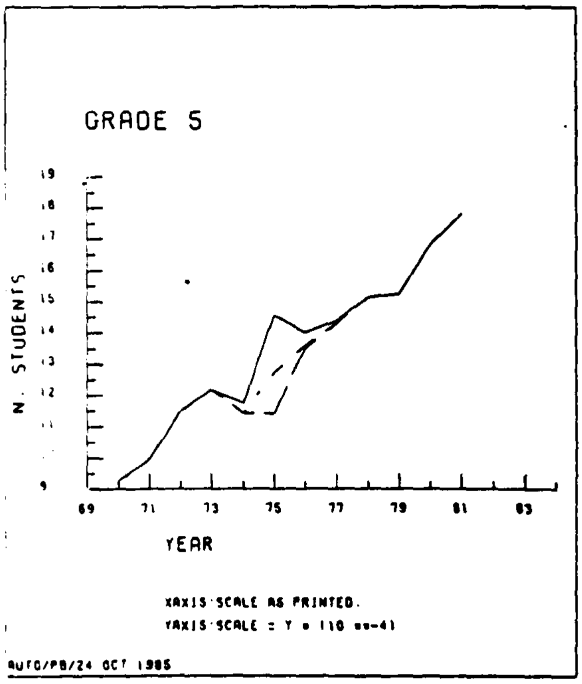
Trans Probab.	p_{55}	p_{56}	p_{66}	p_{67}	p_{77}	p_{78}	p_{88}	p_{89}	p_{99}	$p_{9,10}$	$p_{10,10}$	$p_{10,11}$	$p_{11,11}$	p_{5d}	p_{6d}	p_{7d}	p_{8d}	p_{9d}	$p_{10,d}$	$p_{11,d}$
VAR CODE / YEAR	V1	V2	V3	V4	V5	V6	V7	V8	V9	V10	V11	V12	V13	V14	V15	V16	V17	V18	V19	V20
1971	.16	.67	.17	.63	.23	.67	.26	.58	.21	.58	.04	.77	.11	.17	.19	.10	.16	.21	.23	.89
1972	.15	.73	.16	.66	.22	.68	.22	.64	.23	.64	.06	.74	.12	.12	.18	.10	.14	.12	.20	.88
1973	.13	.75	.16	.79	.25	.64	.20	.66	.23	.60	.06	.72	.14	.12	.06	.11	.14	.17	.22	.86
1974	.06	.84	.03	.89	.13	.78	.14	.95	.10	1.08	.07	.91	.06	.10	.08	.09	.09	-.17	.02	.94
1975	.13	.84	.17	.86	.14	.64	.20	.70	.28	.72	.05	.82	.20	.03	-.03	.22	.10	.00	.13	.80
1976	.10	.77	.13	.71	.10	.90	.19	.77	.28	.68	.10	.96	.24	.13	.16	.00	.04	.04	-.06	.76
1977	.15	.73	.15	.67	.18	.74	.19	.72	.35	.52	.09	.87	.28	.12	.18	.08	.09	.13	.04	.72
1978	.19	.69	.18	.59	.20	.55	.20	.57	.21	.41	.03	.60	.22	.12	.23	.25	.23	.39	.37	.78
1979	.19	.66	.18	.58	.26	.57	.25	.67	.32	.51	.10	.82	.10	.13	.25	.17	.08	.17	.08	.90
1980	.22	.70	.19	.59	.26	.60	.26	.63	.27	.57	.11	.86	.31	.08	.22	.15	.11	.15	.03	.69
1981	.20	.69	.17	.59	.25	.61	.24	.65	.24	.67	.11	.92	.31	.11	.24	.15	.11	.09	-.03	.69
Unr OLS*	.16	.59	.33	.57	.34	.72	.15	1.01	-.12	.06	.95	.81	.23	.25	.10	-.06	-.16	1.06	-.76	.77
Res OLS	.16	.84	.03	.73	.14	.86	.00	.93	.00	.25	.58	.42	.59	.00	.24	.00	.07	.75	.00	.49
Unr GLS*	.19	.43	.52	.81	.01	.43	.39	.78	.11	.55	.06	.49	1.50	.38	-.33	.56	-.17	.34	.45	-.50
Unr.OLS for Matrix A*	.16	.72	.16	.62	.25	.65	.23	.88	.02	-.01	1.09	.72	.37	.12	.22	.10	-.01	.99	-.81	.63
Res.OLS for Matrix A	.18	.59	.33	.66	.22	.79	.09	.91	.00	.41	.24	.76	.27	.23	.00	.00	.00	.59	.00	.73
Res.OLS for Matrix S	.21	.71	.19	.60	.35	.60	.25	.54	.34	.51	.17	.82	.02	.08	.20	.04	.20	.16	.01	.97
Res.OLS for Matrix A1	.18	.58	.35	.65	.21	.79	.09	.91	.00	.54	.00	1.00	.00	.24	.00	.00	.00	.46	.00	1.00

*The drop-out probabilities were calculated by difference, assuming the row-sum condition is satisfied; they are not the OLS estimates although the standard errors are higher in the second case for more than half of the estimates. Furthermore, the estimates obtained in this second case are not so close to the assumed constant probabilities when compared to the estimates obtained for the matrix $\underline{A}$ of school enrolment. This was the reason why the iterative process was stopped after the second run. The confidence intervals still present a wide range, even at the 90% level, which implies that in some cases, the set of the observed point estimates for the period of analysis still does not fall within the range of the 95% confidence interval ($\hat{p}_{56}$, $\hat{p}_{89}$ in the case of the matrix $\underline{A}$ and $\hat{p}_{56}$, $\hat{p}_{66}$, $\hat{p}_{89}$ in the case of matrix $\underline{A1}$). However, the t-values indicate greater reliability for the estimates of the transition probabilities. The F-statistic is practically the same for the two restricted estimation procedures undertaken.

From Table 4.21, which compares the different estimates obtained for the transition probabilities, it is clear that there is little difference between the results presented, so that one cannot infer that one method is preferred in the sense that it gives a closer proxy for the transition probabilities.

Figures 4.1 to 4.7 illustrate the school enrolment per school year and per grade, for the three matrices of enrolment $\underline{N}$, $\underline{A}$ and $\underline{A1}$. What is noteworthy from these graphs is that the attempt to dissociate the returnee students from the observed school enrolment present a similar structure to the observed values for most of the grades. In the case of grades 5, 6 and 9 the iterative process seems to be quite efficient converging the number of students enrolled into a smoother series of values; the same does not occur for the other grades, in which the attempt at dissociating the returnee students

Fig. 4.1-4.7 School Enrolment by Grade - WHOLE COUNTRY



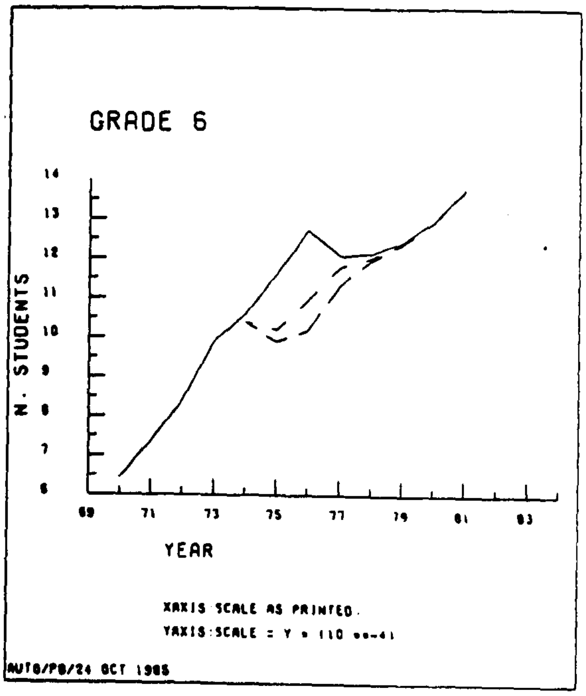
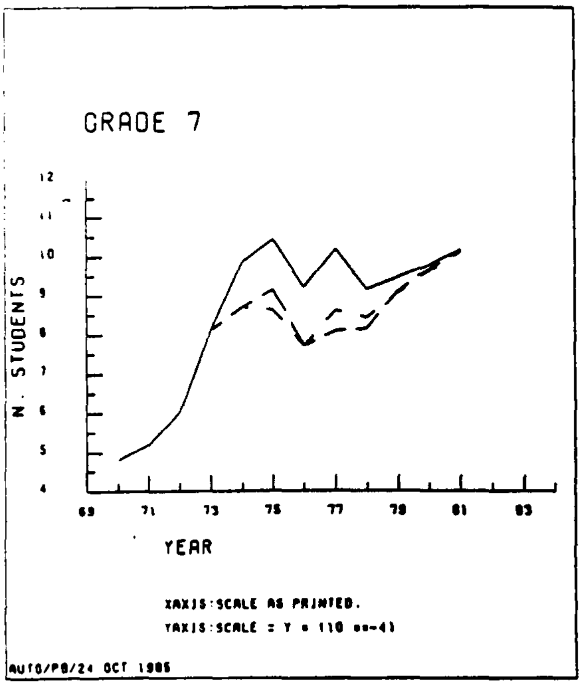


Fig. 4.1-4.7 School Enrolment by Grade - WHOLE COUNTRY
(Continued)



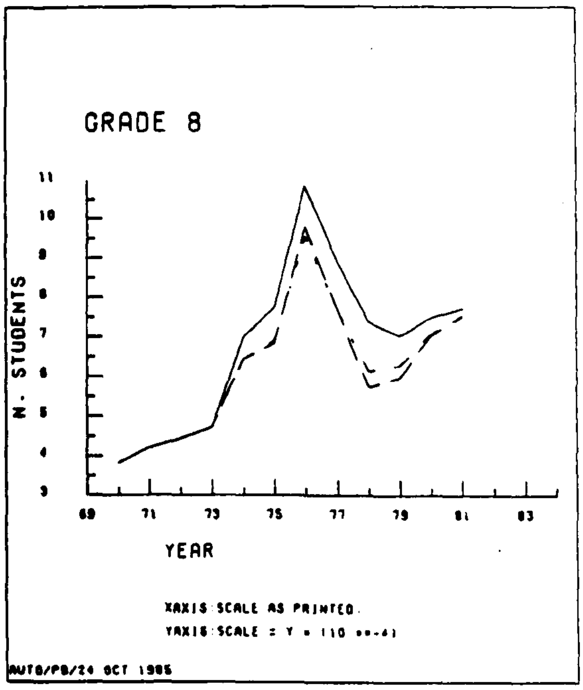
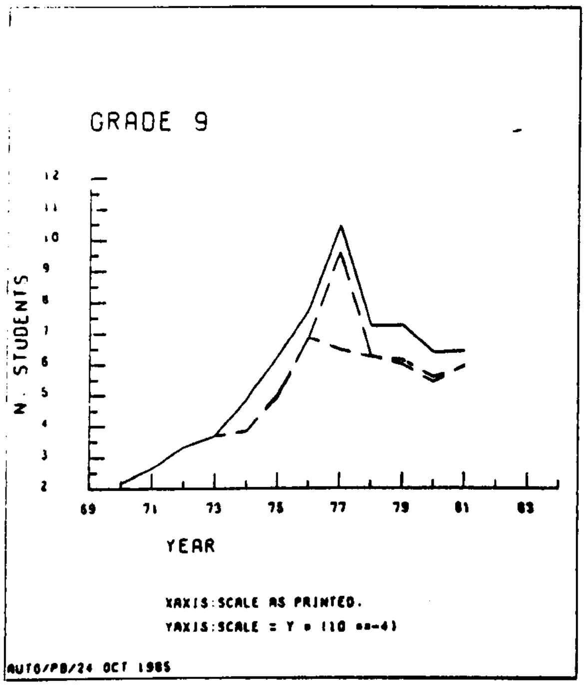
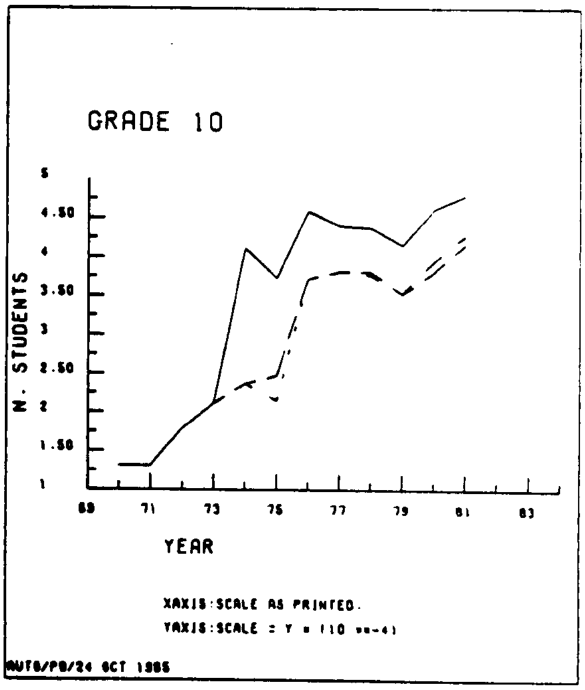
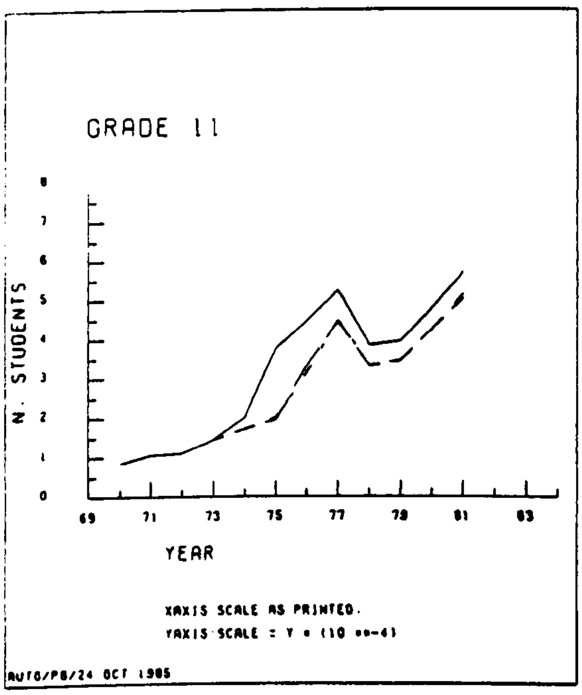


Fig. 4.1-4.7 School Enrolment by Grade - WHOLE COUNTRY
(Continued)







——— Original Data (Matrix N)

— — — Smoothed Data (Matrix A)

-·-·-·- Smoothed Data (Matrix A1)

does not produce significant changes in the series of the enrolment. We can conclude, therefore, that the assumption of constant transition probabilities is quite limited and unreliable when trying to study the behaviour of the school enrolment of the Portuguese educational system, through the corresponding transition probabilities.

Chapter 4

Footnotes

1. This does not include the islands of Madeira and Azores.

2. A summary description of the programs used in this chapter and displayed in Appendix B is presented in Appendix G.

3. The total number of places in higher education, as a whole, as well as in each sector, each institution and each speciality, is fixed by the Ministry of Education.

4. The resulting estimator of the transition probabilities is called the two-stage AITKEN estimator because its value is calculated in two stages. This estimator is asymptotically equivalent to AITKEN's generalised least squares estimator and, therefore, is asymptotically efficient and has a normally asymptotic distribution [see KAKWANI, 1967].

5. A FORTRAN program using a NAG Library routine was also tried but the same results occured.

Chapter 5

REGIONAL APPLICATION OF THE BASIC MARKOV MODEL

5.1. Introduction

A regional level application of the basic Markov model is now performed in order to improve the knowledge of the behaviour of the different point estimates of the transition probabilities and their estimators, as well as to detect regional disparities in their behaviour. The availability of published data, by district justifies the use of the 'district' as the base unit of this analysis.

It is well known that continual aggregation tends to dilute the original information. As soon as a regional level analysis starts, not only global migratory movements can exist but also those that are educational in origin, for example, migration from rural to urban areas in order to attend secondary school.[1] Migration between neighbouring districts can result easily in cases where students living near the frontier of a district complete the preparatory level in a school located in the district and attend a secondary school in a neighbouring district. The problem of population migration is liable to introduce serious distortions into comparisons as soon as it reaches certain proportions. However, due to the lack of sufficient information, the problem of migration will not be taken into account in this analysis. Therefore, some distortions may occur in the data, generating underestimated or overestimated point estimates for the coefficients relating different education levels.

This can be seen in Tables C.6.a to C.6.r in Appendix C, where several point estimates of the transition probabilities are negative or greater then unity (negative drop-outs and promotion rates above unity). These values show that some disturbance took place, as these coefficients should stay in the range 0-1. Because of the necessity of absorbing the nationals from Angola and Mozambique, after these colonies achieved independence, it is understandable to observe some distortions in the point estimates of the probabilities. It can be seen, however, that in some districts, the same phenomenon occurs before 1974, or after the period of the disturbance, which suggests the existence of migratory factors, changes from the private sector to the public sector schools, or errors in the data gathering by the public institutions.

The same structure of analysis developed in Chapter 4 for the whole country, will now be applied to the eighteen districts into which the country is administratively divided. Unrestricted and restricted OLS estimators for the transition probabilities will be computed and an analysis of the influence of the returnees will be carried out. The GLS estimator is not applied in this chapter as the study undertaken for the whole country has shown the restricted OLS estimator to be the one which yields, as we have argued, more accurate estimates of the transition probabilities.

5.2. Regional Aspects of the Education System in Portugal

Portugal is a small country covering an area of 92 500 km^2 (including the islands of Madeira and Azores) with a population of less than ten million. The population is unevenly distributed over the country, the north being more densely populated than the south and the coast more densely populated than the interior. Portugal has the largest percentage of children of the age group 10-14 among OECD countries [OECD, 1984]. The regional age distribution of the population in the seventies presented in Table 5.1 shows that the less populated districts, Beja, C. Branco, Evora and Portalegre, associated with more rural areas, have had a decrease in the population of the two age groups 0-14 and 15-19. The highest increase appears in Braga, Lisboa, Porto and Setubal, that is, the two main districts of the country (Lisboa and Porto) and their satellite industrialised districts (Braga and Setubal). As has already been discussed, about 350 000 of those returning from the former colonies were under sixteen. A concentration of the returnees around the capital and the second town may have contributed to the observed increase in these districts of the youth population.

Table 5.2 Age Distribution of the Population in Percentage

Year	Age group 10-14	 15-25
1971	28.4	16.5
1981	26.0	16.6

An increase in the youth population should imply an increase in school enrolment. Table 5.3 presents the increase of school enrolment during the 1970s, by district and by level of education, in the public sector schools. The districts with a decrease in the youth population of age 10-19 (Beja, C. Branco, Evora and Portalegre) reveal, as expected, a lower increase in the school enrolment in the basic preparatory and unified general course. Nevertheless, the same does not occur for the districts with the highest increase in the youth population (Braga, Lisboa, Porto and Setubal). These

Table 5.1 Youth Population by Age Group 10-14 and 15-19

Year	1970		1981		1981/1970	
Age Groups	10-14	15-19	10-14	15-19	10-14	15-19
Whole Country*	753920	681910	803833	808508	1.1	1.2
Districts						
Aveiro	57125	49695	59733	60910	1.0	1.2
Beja	17675	15665	14162	14238	0.8	0.9
Braga	71375	60870	78819	78148	1.1	1.3
Braganca	20825	17150	17893	17981	0.9	1.0
C.Branco	23420	22630	17446	19223	0.7	0.9
Coimbra	34370	32240	34628	34954	1.0	1.1
Evora	14170	13780	12626	13328	0.9	1.0
Faro	19890	20350	22195	22329	1.1	1.1
Guarda	21505	17875	16939	18103	0.8	1.0
Leiria	35775	33435	36252	37147	1.0	1.1
Lisboa	112985	112520	153469	151558	1.4	1.4
Portalegre	11445	10785	9741	10361	0.9	1.0
Porto	138845	124030	151033	155415	1.1	1.3
Santarem	36075	33645	34935	35880	1.0	1.1
Setubal	37205	34790	49471	47064	1.3	1.4
C.Castelo	26205	21580	24946	23935	0.9	1.1
V.Real	31260	25350	28154	27928	0.9	1.1
Viseu	43770	35520	41391	40006	0.9	1.1

SOURCE: INE, XI, XII Census *Without islands

Table 5.3 School Enrolment by Level of Education
(Public Sector Schools)

Level	Basic Preparatory			Secondary Unified General Course			Secondary Complementary Course		
Year	1970/71	1981/82	81/82 / 70/71	1970/71	1981/82	81/82 / 70/71	1970/71	1981/82*	81/82 / 70/71
Whole Country	256914	298041	1.9	108014	236439	2.2	21553	94113	4.4
Districts									
Aveiro	12722	23135	1.8	5677	13792	2.4	679	5448	8.0
Beja	3465	5489	1.6	1993	4578	2.3	193	1486	7.7
Braga	12147	25808	2.1	5352	12616	2.4	967	5633	5.8
Braganca	3621	7516	2.1	2449	6026	2.5	437	2337	5.3
C.Branco	4411	7342	1.6	3405	6169	1.8	576	2998	5.2
Coimbra	7270	14965	2.1	5921	11797	2.0	1436	6164	4.3
Evora	3518	5202	1.5	2389	5092	2.1	495	2452	5.0
Faro	5243	9968	1.9	5140	9291	1.8	1028	3698	3.6
Guarda	2467	6485	2.6	1869	4320	2.3	435	2572	5.9
Leiria	5968	14306	2.4	3518	9090	2.6	621	3191	5.1
Lisboa	32411	62795	1.9	31537	64268	2.0	7068	29537	4.2
Portalegre	2439	4103	1.7	1622	3244	2.0	289	1737	6.0
Porto	26496	56630	2.1	17301	37518	2.2	3928	15972	4.2
Santarem	7696	14006	1.8	4909	11294	2.3	857	5042	5.9
Setubal	10283	23563	2.3	8060	22841	2.8	908	7875	8.7
V.Castelo	4802	9193	1.9	1416	4737	3.3	336	2072	6.2
V.Real	4803	9920	2.1	2865	7816	2.7	561	2986	5.3
Viseu	7152	15601	2.2	2491	8979	3.6	727	3864	5.3

* Does not include 12th year

SOURCE: Diagnostico/Previsoes, Educational Planning Bureau (GEP), Ministry of Education, 1983 districts, with the exception of Setubal, do not show, by contrast, the highest increase in the school enrolment. It seems reasonable to infer that some of the young returnees within the schooling age, did not attend public sector schools.[2]

The rural sector is large in Portugal, with a poor infrastructure of roads and large, thinly populated tracts. The network of provisions for preparatory education is inefficient. The bussing of children to distant schools, designed to allow equal opportunities in all areas to meet educational needs, has been one measure adopted. However, wholesale bussing is not an easy solution; in some cases boarding away from home is necessary. There is also an integrated monitor - television instructional system CPTV (ciclo preparatorio-televisao), which serves 58 000 children in 1 150 'posts', not only located in the remoter areas but also in towns and suburban areas. CPTV is playing a significant role and has been an essential instrument for reducing the drop-outs in the more inaccessible parts of the country, and in improving the attendance at compulsory schooling.

The percentages of students who leave the school system during or after preparatory schooling are presented, by district, in Table 5.4. The overall distribution of the drop-out rates for preparatory and secondary levels of education can be seen also in Table 5.4 and in Table 5.5. It is apparent that the drop-out rates have decreased between 1971 and 1981 in all districts, with the exception of Lisboa, Setubal and Faro. These districts show a significant increase in the drop-out rates. However, the absolute values of their drop-out rates in the preparatory level are the lowest observed in all the country (less than 12%). Braga and V. do Castelo, in the North of the

Table 5.4 Percentage of Dropouts in Preparatory and Secondary Schooling, 1971-1981

Level	Basic Preparatory		Basic Preparatory and Secondary	
	Year		Year	
District	1971	1981	1971	1981
Aveiro	26.7	21.4	22.7	20.9
Beja	26.6	15.3	23.1	15.8
Braga	30.7	28.6	26.9	25.3
Braganca	16.2	14.4	16.7	16.2
C.Branco	19.6	15.0	22.2	18.6
Coimbra	12.7	13.9	14.3	13.8
Evora	19.9	12.3	20.0	12.2
Faro	7.2	11.3	12.9	14.1
Guarda	19.8	18.4	20.7	15.8
Leiria	23.7	19.7	24.6	19.7
Lisboa	2.6	10.2	11.7	14.6
Portalegre	20.1	16.9	20.4	15.8
Porto	20.0	18.4	22.2	19.0
Santarem	20.2	15.2	20.6	17.1
Setubal	3.5	9.1	9.9	16.9
V.Castelo	39.8	27.1	34.4	22.6
V.Real	30.6	15.4	28.3	15.9
Viseu	34.3	22.7	30.3	20.4

SOURCE: Diagnostico/Previsoes, Educational Planning Bureau (GEP), Ministry of Education, 1983

Table 5.5 District Distribution of the Overall Dropouts in Public Schools

(Basic Preparatory and Secondary Levels of Education)

District	1971	1972	1973	1974	1975	Year 1976	1977	1978	1979	1980	1981
Aveiro	8.0	9.4	8.1	12.4	7.7	9.4	10.9	7.1	8.1	8.2	7.6
Beja	2.4	2.6	3.2	4.9	1.6	2.9	1.8	1.6	2.3	2.6	1.6
Braga	9.6	9.2	10.2	11.0	8.4	12.8	8.3	8.0	9.0	7.3	9.9
Braganca	2.0	2.1	2.3	5.3	0.7	3.6	2.8	1.2	3.0	2.5	2.2
C.Branco	3.0	2.5	4.9	4.6	2.2	4.1	2.9	1.9	2.1	2.3	2.8
Coimbra	3.9	6.2	5.1	3.1	7.8	5.3	4.8	4.0	5.1	5.0	3.9
Evora	2.4	1.7	2.5	4.4	3.7	-9.8	2.2	1.9	2.1	1.1	1.3
Faro	2.7	3.4	3.6	2.2	0.9	2.7	2.9	3.9	2.5	3.6	2.8
Guarda	1.8	0.5	2.2	4.2	-3.0	4.0	2.7	1.8	2.3	2.3	1.8
Leiria	4.6	2.9	2.7	4.5	7.1	4.8	5.0	4.7	5.2	4.8	4.5
Lisboa	15.5	17.0	19.6	-7.0	26.6	19.4	11.9	23.4	19.0	19.7	19.9
Portalegre	1.6	1.7	2.5	3.4	0.6	2.0	1.7	1.0	1.1	1.4	1.2
Porto	19.6	16.8	12.6	15.4	22.7	15.6	17.8	16.5	16.6	18.1	17.8
Santarem	5.1	5.7	2.8	6.5	1.1	8.3	8.5	5.5	4.2	4.4	4.5
Setubal	3.5	6.4	2.8	-1.2	7.6	1.6	3.6	8.2	6.8	5.3	7.9
V.Castelo	4.2	3.5	3.8	6.1	7.0	1.6	4.0	2.7	2.5	3.5	3.0
V.Real	4.3	3.5	3.9	6.7	-5.0	5.4	3.5	2.1	3.5	3.1	2.4
Viseu	5.8	4.9	7.2	13.5	2.3	6.3	4.7	4.5	4.6	4.8	4.9
TOTAL	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0	100.0

SOURCE: Diagnostico/Previsoes, Educational Planning Bureau (GEP), Ministry of Education, 1983.

country, present the highest drop-out rates, both at this level of education (over 27%), and in both preparatory and secondary levels (over 22%).

It is also important to note the inequalities between interior and coastal districts and between rural and urban areas, which become significant in upper secondary education, where it is difficult to equalise opportunities of access for every population settlement. The network of the lower secondary level is larger than that for the upper secondary level and the five branches ('areas') of study are not available in all schools. Furthermore, from the 32 vocational specialisms within the secondary system, only between one and six appear in any one school, which sometimes leads to a lack of interest in the specialisms offered by the school.

With respect to the observed promotion point estimates between the preparatory level and the lower secondary level (see Tables C.6.a - C.6.r in Appendix C), it can be seen that in V. do Castelo and Viseu, the coefficients have not reached 50%, which means that in these two districts less than half of the young people enrolled in the last grade of the compulsory schooling intended to pursue their studies. On the other hand, Lisboa, Faro and Setubal are the districts with higher promotion rates[3] between these two levels of education.

It is also interesting to note that the districts which present a large shortfall in preparatory schooling, have, as expected, successively lower drop-out rates in secondary schooling. Beja and Braganca are the districts which present the lowest promotion rates, with high drop-out rates between the secondary unified general course and the secondary complementary course.

The number of students enrolled in each grade, for the period of analysis and by district, are presented in Appendix C, Tables C.1.a - C.1.r.

5.3. The OLS Estimator

Unrestricted and restricted least squares estimates were obtained for all districts, using the same method discussed in the preceding chapter. The multiple regression program included in the SPSS package and the quadratic programming algorithm SYMQUAD included in the MPOS system of programs, were applied and the results are given in Table 5.6.

The disturbances observed in the different time-series for the whole country still hold in all the districts individually. The unrestricted OLS transition probability estimates are poor, with the non-negativity and the row-sum conditions being violated. In particular, almost all the drop-out probability estimates lie outside the range 0-1, which has already been observed for the whole country estimates. When the row-sum and non-negativity conditions are introduced, the restricted OLS estimators computed yield estimates a little closer to the point estimate parameters. However, in some cases the values obtained are completely out of the range of acceptance, taking into account the series of observed values for the point estimates. Moreover, the estimates of the drop-out probabilities obtained by the restricted OLS procedure are almost all zero. Several attempts have been made to avoid this problem, delimiting the bounds for the minimisation problem in order to force

Table 5.6 The OLS Estimates of the Transition Probabilities

District	Aveiro		Beja		Braga		Braganca		C.Branco		Coimbra	
Method	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS
p_{55}	.15	.16	.15	.14	.16	.13	.12	.13	.15	.15	.16	.16
p_{56}	.56	.53	.76	.63	.34	.46	.72	.69	.30	.53	.53	.63
p_{66}	.33	.37	.12	.27	.59	.43	.17	.23	.69	.42	.44	.31
p_{67}	.43	.67	1.01	.73	.35	.57	1.10	.77	.35	.58	.45	.69
p_{77}	.41	.15	-.39	.00	.48	.17	-.31	.12	.54	.21	.52	.21
p_{78}	.80	.85	.59	.71	.66	.83	.67	.88	.68	.79	.68	.79
p_{88}	.06	.00	.29	.11	.20	.00	.23	.00	.28	.14	.22	.08
p_{89}	1.02	.84	.96	.89	.85	.85	1.01	1.00	.51	.52	.97	.92
p_{99}	-.22	.00	-.04	.00	-.02	.00	.08	.06	.47	.45	-.04	.00
$p_{9,10}$	.19	.41	.20	.23	-.05	.40	.18	.42	.44	.25	.05	.19
$p_{10,10}$	.70	.20	.60	.39	1.13	.32	.63	.02	.33	.62	1.01	.72
$p_{10,11}$	.81	.80	.90	.61	.97	.68	.58	.07	.48	.38	.74	.28
$p_{11,11}$	.29	.17	.23	1.00	.07	.28	.52	1.00	.64	.74	.34	.78
p_{5d}	.31	.31	1.45	.23	1.86	.40	.04	.18	.97	.32	.60	.21
p_{6d}	.61	.00	-1.68	.00	-3.33	.00	.26	.00	.80	.00	.13	.00
p_{7d}	-1.35	.00	.51	.29	.88	.00	.31	.00	-.99	.00	-1.58	.00
p_{8d}	.13	.16	-.01	.00	-.06	.15	.74	.00	1.15	.34	-.12	.00
p_{9d}	2.03	.59	1.63	.77	1.83	.60	.09	.52	.74	.30	2.17	.81
$p_{10,d}$	-1.08	.00	-1.49	.00	.49	.00	-4.14	.91	.07	.00	1.52	.00
$p_{11,d}$	.63	.83	-.13	.00	1.66	.72	3.77	.00	.15	.26	-1.53	.22

Table 5.6 **The OLS Estimates of the Transition Probabilities**
(continued)

District	Evora		Faro		Guarda		Leiria		Lisboa		Portalegre	
Method	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS
p_{55}	.16	.18	.17	.18	.14	.13	.13	.18	.17	.18	.15	.15
p_{56}	.84	.66	.59	.82	.83	.87	.61	.66	.77	.82	.52	.51
p_{66}	.07	.29	.38	.13	.06	.00	.27	.30	.17	.12	.44	.44
p_{67}	.32	.72	.36	.74	.58	.33	.51	.65	.53	.88	.63	.56
p_{77}	.64	.14	.64	.21	.14	.29	.18	.16	.48	.12	.09	.20
p_{78}	.85	.86	.58	.79	.85	.71	.79	.84	.56	.78	.78	.80
p_{88}	.02	.00	.33	.08	.09	.00	.13	.06	.36	.10	.11	.09
p_{89}	.83	.92	.95	.92	1.24	1.00	1.37	.94	.99	.90	.67	.71
p_{99}	.19	.07	-.02	.00	-.18	.20	-.37	.00	-.11	.00	.32	.28
$p_{9,10}$	.58	.33	.12	.50	.33	.80	.13	.00	-.02	.34	.24	.44
$p_{10,10}$	.05	.38	.84	.15	.54	.00	.85	.48	.09	.47	.71	.37
$p_{10,11}$	.32	.36	.94	.84	.87	1.00	1.08	.52	.85	.53	.75	.63
$p_{11,11}$	.79	.72	.01	.07	.23	.00	-.01	.00	.22	.49	.36	.49
p_{5d}	.53	.17	.59	.00	.70	.00	.64	.17	8.05	.00	.90	.34
p_{6d}	-.25	.00	2.50	.13	-.45	.67	.87	1.05	-7.02	.00	-.58	.00
p_{7d}	-2.52	.00	-1.18	.00	-1.13	.00	-1.30	.00	-4.55	.10	-.58	.00
p_{8d}	-.12	.00	2.16	.08	-.66	.00	.18	.00	-.06	.00	-4.06	.00
p_{9d}	2.07	.60	2.89	.50	1.81	.00	2.86	1.00	4.87	.66	.71	.28
$p_{10,d}$	.75	.27	-3.90	.00	-.69	.00	1.39	.00	11.56	.00	-2.21	.00
$p_{11,d}$	-1.90	.28	.29	.93	.83	1.00	-1.59	1.00	-5.64	.51	1.44	.51

Table 5.6 **The OLS Estimates of the Transition Probabilities**
(continued)

District	Porto		Santarem		Setubal		V.Castelo		V.Real		Viseu	
Method	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS
p_{55}	.17	.18	.15	.16	.20	.20	.13	.14	.15	.13	.14	.14
p_{56}	.54	.61	.60	.84	.30	.76	.35	.42	.28	.55	.88	.48
p_{66}	.38	.29	.34	.05	.72	.18	.60	.52	.69	.35	-.03	.44
p_{67}	.50	.71	.79	.72	.39	.82	.38	.48	.34	.65	.36	.56
p_{77}	.44	.18	.09	.16	.63	.20	.38	.22	.65	.25	.61	.10
p_{78}	.77	.82	.65	.84	.83	.81	.94	.78	.81	.75	.65	.90
p_{88}	.04	.00	.23	.00	.00	.04	-.25	.00	-.02	.09	.22	.00
p_{89}	.99	.90	.96	.86	1.12	.96	.82	.86	.87	.91	1.08	1.00
p_{99}	-.11	.00	-.04	.00	-.19	.00	.12	.02	.08	.04	-.25	.00
$p_{9,10}$	.20	.24	-.10	.30	-.02	.02	.29	.35	.35	.41	.05	.17
$p_{10,10}$	.69	.54	1.23	.40	1.09	.93	.43	.15	.38	.25	.94	.71
$p_{10,11}$	.58	.46	1.18	.60	.87	.06	.62	.59	1.17	.75	.87	.29
$p_{11,11}$	.48	.53	-.25	.25	.16	.97	.53	.39	-.18	.32	.21	.76
p_{5d}	3.03	.21	.59	.00	4.95	.03	4.12	.44	.41	.32	.64	.38
p_{6d}	-4.79	.00	-.88	.23	-6.27	.00	-8.21	.00	.73	.00	.76	.00
p_{7d}	-1.04	.00	-.49	.00	-1.37	.00	2.79	.00	-2.16	.00	-2.74	.00
p_{8d}	1.79	.10	2.25	.14	.48	.00	2.47	.14	1.01	.00	.60	.00
p_{9d}	2.06	.76	1.04	.70	2.67	.98	2.62	.63	1.25	.55	2.39	.83
$p_{10,d}$	1.99	.00	-3.21	.00	3.77	.00	3.75	.26	-1.43	.00	-1.30	.00
$p_{11,d}$	-1.02	.47	1.80	.75	-2.63	.03	-1.44	.61	1.77	.68	1.73	.24

the probabilities to fall into a more acceptable range. No reasonable estimates have been found and very high opportunity costs result for those probabilities that had distorted estimators.

The statistics for the drop-out probability estimates are not studied in this chapter because of the unreal values obtained for these estimates. At this stage attention will be paid to the analysis of the consistency of the repetition and promotion probability estimates. FORTRAN program RESIDD applied to each district individually gives the statistics for the restricted OLS estimates.

The reliability of the partial regression estimates, for EQ1 to EQ7 of equations (4.1) can be examined using the t-values for the unrestricted and the restricted OLS estimates of the transition probabilities presented in Table 5.7. Most of the equations show that only one of the independent variables has statistical significance. However, this significance is higher in the case of the restricted OLS estimator. These results are similar to those obtained for the whole country, which is not unexpected due to the high degree of multicollinearity between the independent variables. Furthermore, and for some of the districts, low t-values for both probability estimates result when the equation relates two different levels of education.

Globally, the 95% confidence intervals for the restricted OLS estimates presented in Table 5.9 are smaller than those for the unrestricted OLS estimates presented in Table 5.8. Table 5.9 shows that in only four districts (Evora, Faro, Guarda and Leiria) the 95% confidence intervals for the set of the restricted OLS transition

Table 5.7 T-values for the OLS Estimates of the Transition Probabilities

District	Aveiro		Beja		Braga		Braganca		C.Branco		Coimbra		Evora		Faro		Guarda	
Method	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS
p_{55}	8.82	1.38	9.37	1.22	10.67	1.36	10.01	1.59	10.72	1.48	8.89	1.69	8.42	0.81	10.02	1.88	10.01	1.83
p_{56}	4.24	5.83	4.00	6.53	1.99	4.18	4.80	5.48	1.30	3.79	2.44	7.35	2.57	2.67	2.50	6.89	3.74	10.88
p_{66}	1.92	4.38	0.53	2.73	2.42	4.66	0.91	2.11	2.57	2.85	1.70	4.07	0.19	0.65	1.41	1.10	0.22	0.00
p_{67}	1.95	3.30	3.45	4.08	1.61	2.73	4.11	4.40	1.48	3.28	1.42	3.85	1.80	4.00	1.19	3.54	2.29	0.98
p_{77}	1.26	0.77	0.95	0.00	1.41	1.04	2.42	0.58	1.58	1.58	1.38	1.33	2.77	0.76	1.88	1.19	0.35	1.39
p_{78}	4.97	6.40	3.39	4.06	4.13	6.01	4.00	5.95	3.11	4.47	3.82	5.53	4.15	4.94	3.37	4.44	3.99	2.42
p_{88}	0.30	0.00	1.29	0.48	0.99	0.00	1.47	0.00	1.15	0.66	1.03	0.52	0.08	0.00	1.59	0.41	0.38	0.00
p_{89}	7.91	9.45	7.74	9.89	6.20	9.20	5.34	9.52	3.89	4.60	7.89	11.08	3.37	6.57	4.30	7.49	5.66	6.25
p_{99}	1.40	0.00	0.40	0.00	0.12	0.00	0.45	0.38	3.22	3.46	0.01	0.00	0.75	0.46	0.08	0.00	0.84	0.94
$p_{9,10}$	1.26	1.81	2.22	1.01	0.26	4.68	1.49	2.04	2.64	1.53	0.32	0.95	2.48	1.21	0.53	1.95	2.02	2.81
$p_{10,10}$	2.15	0.88	2.75	4.68	3.60	2.94	2.33	0.23	1.15	5.90	3.95	8.58	0.11	2.48	2.42	1.01	2.06	0.00
$p_{10,11}$	4.45	6.70	5.66	1.72	3.44	4.69	1.76	0.40	3.08	4.69	4.38	2.29	0.89	1.71	4.12	4.69	4.71	7.84
$p_{11,11}$	1.60	1.09	1.40	2.16	0.23	1.68	1.41	5.08	3.79	7.64	1.89	5.20	2.07	2.60	0.04	0.27	1.20	0.00

$*t^{9}_{0.025} = 2.262; \; t^{8}_{0.025} = 2.306$

Table 5.7 T-values for the OLS Estimates of the Transition Probabilities (continued)

District	Leiria		Lisboa		Portalegre		Porto		Santarem		Setubal		V.Castelo		V.Real		Viseu	
Method	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS
p_{55}	8.67	2.00	15.34	2.72	10.03	1.33	10.02	1.88	10.72	1.55	12.50	2.60	8.67	1.31	3.83	1.34	8.75	1.37
p_{56}	3.51	6.78	3.44	1.07	2.29	3.59	4.26	7.01	2.71	7.09	0.97	6.55	1.93	3.44	0.72	2.59	6.52	4.80
p_{66}	1.24	0.33	1.70	1.54	1.67	3.02	2.10	4.33	1.29	0.40	1.98	1.70	2.52	3.71	1.37	1.91	0.17	4.11
p_{67}	2.02	2.90	1.48	4.33	4.53	4.96	1.69	4.06	2.37	3.96	0.76	5.20	2.88	3.86	1.63	4.89	4.09	7.95
p_{77}	0.49	0.95	1.27	0.67	0.43	1.90	1.23	1.05	0.23	0.72	1.17	1.26	1.60	2.01	2.44	1.68	3.94	1.73
p_{78}	4.73	4.91	2.21	4.02	5.34	6.91	3.58	4.56	4.64	6.68	3.97	5.03	4.33	4.13	4.33	7.08	2.64	6.62
p_{88}	0.62	0.36	1.18	0.49	0.63	0.93	0.15	0.00	1.31	0.00	0.00	0.23	0.85	0.00	0.08	0.97	0.76	0.00
p_{89}	9.32	6.98	5.50	8.11	5.60	0.88	8.84	10.22	7.27	6.70	0.55	7.94	3.12	5.77	3.58	9.38	6.75	10.31
p_{99}	2.13	0.00	0.52	0.00	2.42	2.50	0.85	0.00	0.27	0.00	0.85	0.00	0.40	0.13	0.30	0.37	1.53	0.00
$p_{9,10}$	0.67	0.00	0.07	1.58	1.52	2.88	1.12	1.56	0.52	1.09	0.11	0.11	1.46	0.88	2.32	2.68	0.56	1.24
$p_{10,10}$	2.46	1.87	0.19	3.48	2.74	3.43	2.30	4.03	3.62	2.62	3.10	10.45	0.95	0.95	1.30	3.09	5.08	11.20
$p_{10,11}$	3.97	1.42	3.26	3.73	4.52	6.49	3.24	3.54	5.90	4.03	2.81	0.30	1.80	1.93	2.96	4.60	3.64	1.87
$p_{11,11}$	0.04	0.00	0.82	2.72	2.02	4.08	2.45	3.70	1.08	1.57	0.48	4.04	1.46	1.13	0.39	1.90	0.86	4.28

Table 5.8 95% Confidence Intervals for the Unrestricted OLS Estimates of the Transition Probabilities

District	Aveiro		Beja		Braga		Braganca		C.Branco		Coimbra	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	0.112	0.187	0.114	0.187	0.131	0.199	0.087	0.142	0.112	0.176	0.122	0.202
p_{56}	0.262	0.860	0.331	1.191	-0.050	0.726	0.379	1.056	-0.224	0.822	0.357	1.017
p_{66}	-0.063	0.715	-0.395	0.640	0.035	1.140	-0.253	0.597	0.084	1.303	-0.150	1.021
p_{67}	-0.068	0.926	0.354	1.678	-0.139	0.842	0.491	1.706	-0.184	0.887	-0.271	1.162
p_{77}	-0.329	1.143	-1.315	0.544	-0.289	1.255	-1.064	0.444	-0.235	1.315	-0.339	1.368
p_{78}	0.432	1.162	0.199	0.985	0.296	1.019	0.375	0.955	0.181	1.174	0.279	1.084
p_{88}	-0.389	0.503	-0.219	0.793	-0.261	0.654	-0.120	0.587	-0.274	0.825	-0.263	0.701
p_{89}	0.731	1.311	0.677	1.239	0.536	1.155	0.578	1.435	0.218	0.811	0.691	1.249
p_{99}	-0.578	0.133	-0.353	0.281	-0.404	0.366	-0.326	0.484	0.143	0.805	-0.346	0.270
$p_{9,10}$	-0.219	0.602	-0.006	0.398	-0.477	0.382	-0.092	0.455	0.058	0.813	-0.308	0.404
$p_{10,10}$	-0.035	1.442	0.102	1.090	0.421	1.840	0.016	1.239	0.321	0.978	0.410	1.614
$p_{10,11}$	0.396	1.221	0.542	1.261	0.331	1.607	-0.167	1.319	0.126	0.833	0.355	1.118
$p_{11,11}$	-0.119	0.701	-0.145	0.597	-0.619	0.760	-0.317	1.353	0.254	1.020	0.070	0.744

Table 5.8 95% Confidence Intervals for the Unrestricted OLS Estimates of the Transition Probabilities (continued)

District	Evora		Faro		Guarda		Leiria		Lisboa		Portalegre	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	0.121	0.205	0.134	0.209	0.104	0.169	0.105	0.172	0.144	0.194	0.116	0.184
p_{56}	0.096	1.579	0.059	1.129	0.330	1.334	0.486	1.277	0.262	1.275	0.002	1.031
p_{66}	-0.779	0.917	-0.230	0.985	-0.555	0.664	-0.518	0.463	-0.419	0.752	-0.153	1.040
p_{67}	-0.081	0.725	-0.325	1.042	0.005	1.148	-0.273	0.866	-0.283	1.250	0.316	0.945
p_{77}	0.119	1.162	-0.137	1.406	-0.762	1.032	-0.227	1.452	-0.385	1.341	-0.384	0.563
p_{78}	0.377	1.325	0.194	0.971	0.373	1.335	0.274	1.030	-0.009	1.125	0.453	1.112
p_{88}	0.550	0.581	-0.141	0.800	-0.446	0.620	-0.250	0.696	-0.336	1.047	-0.291	0.500
p_{89}	0.269	1.384	0.446	1.448	-0.747	1.739	0.743	1.408	0.584	1.399	0.398	0.940
p_{99}	-0.375	0.763	-0.574	0.541	0.660	0.301	-0.646	0.142	-0.585	0.364	0.019	0.616
$p_{9,10}$	0.053	1.111	-0.387	0.631	-0.042	0.694	-0.390	0.488	-0.711	0.668	-0.122	0.594
$p_{10,10}$	-1.049	0.942	-0.007	1.683	-0.061	1.132	0.156	1.720	0.036	2.141	0.125	1.298
$p_{10,11}$	-0.497	1.131	0.450	1.479	0.456	1.292	0.253	1.485	0.261	1.441	0.375	1.126
$p_{11,11}$	-0.075	1.647	-0.563	0.575	-0.209	0.662	-0.407	0.817	-0.388	0.828	0.044	0.763

Table 5.8 95% Confidence Intervals for the Unrestricted OLS Estimates of the Transition Probabilities (continued)

District	Porto		Santarem		Setubal		V.Castelo		V.Real		Viseu	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	0.130	0.206	0.120	0.184	0.161	0.234	0.095	0.160	0.111	0.186	0.098	0.169
p_{56}	0.219	0.856	0.095	1.097	-0.400	0.994	-0.063	0.759	-0.595	1.160	0.301	0.912
p_{66}	-0.032	0.788	-0.257	0.936	-0.100	1.542	0.067	1.142	-0.455	1.830	-0.140	0.677
p_{67}	-0.170	1.165	0.032	1.539	-0.773	1.549	0.080	0.677	-0.136	0.810	0.306	0.704
p_{77}	-0.370	1.245	-0.812	0.984	-0.589	1.854	-0.160	0.914	0.045	1.248	-0.168	0.533
p_{78}	0.293	1.265	0.331	0.964	0.361	1.305	0.455	1.438	0.385	1.231	0.228	1.341
p_{88}	-0.573	0.655	-0.168	0.626	-0.596	0.595	-0.913	0.412	-0.523	0.559	-0.531	0.782
p_{89}	0.732	1.241	0.657	1.254	0.659	1.578	0.219	1.412	0.321	1.421	1.005	1.728
p_{99}	-0.398	0.186	-0.380	0.297	-0.693	0.317	-0.549	0.795	-0.519	0.688	-0.734	0.004
$p_{9,10}$	-0.203	0.604	-0.536	0.334	-0.431	0.384	-0.161	0.737	0.008	0.691	-0.097	0.351
$p_{10,10}$	0.009	1.366	0.465	1.999	0.298	1.890	-0.603	1.456	-0.286	1.038	0.430	1.269
$p_{10,11}$	0.174	0.986	0.726	1.637	0.167	1.570	-0.161	1.401	0.278	2.071	0.544	1.627
$p_{11,11}$	0.035	0.924	-0.770	0.275	-0.600	0.918	-0.291	1.349	-1.229	0.870	-0.555	0.554

Table 5.9 95% Confidence Intervals for the Restricted OLS Estimates of the Transition Probabilities

District	Aveiro		Beja		Braga		Braganca		C.Branco		Coimbra	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	-0.106	0.418	-0.119	0.403	-0.084	0.349	-0.052	0.319	-0.084	0.375	-0.053	0.377
p_{56}	0.322	0.741	0.405	0.850	0.209	0.717	0.398	0.980	0.209	0.852	0.433	0.828
p_{66}	0.164	0.575	0.043	0.500	0.217	0.643	-0.022	0.480	0.081	0.762	0.131	0.483
p_{67}	0.150	1.111	0.316	1.141	0.089	1.051	0.366	1.176	0.171	0.986	0.279	1.107
p_{77}	-0.306	0.600	-0.344	0.344	-0.210	0.544	-0.358	0.595	-0.100	0.514	-0.152	0.577
p_{78}	0.546	1.159	0.302	1.110	0.514	1.152	0.542	1.222	0.385	1.200	0.458	1.117
p_{88}	-0.351	0.351	-0.413	0.638	-0.341	0.341	-0.323	0.323	-0.352	0.627	-0.276	0.436
p_{89}	0.629	1.039	0.678	1.097	0.634	1.059	0.756	1.244	0.262	0.784	0.729	1.112
p_{99}	-0.195	0.195	-0.241	0.241	-0.197	0.197	-0.307	0.424	0.151	0.752	-0.226	0.226
$p_{9,10}$	-0.113	0.934	-0.298	0.754	-0.045	0.838	-0.050	0.894	-0.124	0.631	-0.272	0.642
$p_{10,10}$	-0.078	0.473	0.194	0.579	0.063	0.569	-0.181	0.217	0.382	0.866	0.477	0.972
$p_{10,11}$	0.529	1.076	-0.204	1.431	0.350	1.018	-0.334	0.471	0.189	0.563	-0.007	0.558
$p_{11,11}$	-0.192	0.530	-0.067	2.067	-0.101	0.669	0.456	1.454	0.517	0.964	0.435	1.127

Table 5.9 95% Confidence Intervals for the Restricted OLS Estimates of the Transition Probabilities (continued)

District	Evora		Faro		Guarda		Leiria		Lisboa		Portalegre	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	-0.329	0.682	-0.036	0.398	-0.033	0.289	-0.026	0.381	0.030	0.330	-0.102	0.406
p_{56}	0.089	1.227	0.544	1.094	0.687	1.057	0.433	0.882	0.641	0.998	0.185	0.840
p_{66}	-0.746	1.317	-0.141	0.401	-0.196	0.196	0.091	0.506	-0.063	0.299	0.109	0.780
p_{67}	0.298	1.131	0.255	1.200	-0.454	1.112	0.135	1.169	0.413	1.351	0.295	0.816
p_{77}	-0.291	0.564	-0.199	0.615	-0.185	0.773	-0.230	0.543	-0.295	0.530	-0.038	0.444
p_{78}	0.463	1.264	0.382	1.202	0.030	1.382	0.449	1.237	0.335	1.228	0.530	1.106
p_{88}	-0.411	0.411	-0.372	0.525	-0.754	0.754	-0.325	0.447	-0.377	0.573	-0.129	0.318
p_{89}	0.594	1.247	0.641	1.207	0.630	1.370	0.656	1.222	0.645	1.159	0.519	0.892
p_{99}	-0.380	0.523	-0.323	0.323	-0.290	0.695	-0.310	0.310	-0.317	0.317	0.021	0.536
$p_{9,10}$	-0.296	0.959	-0.930	1.094	0.138	1.456	-1.174	1.174	-0.155	0.842	0.088	0.793
$p_{10,10}$	0.023	0.728	-0.191	0.497	-0.402	0.402	-0.115	1.068	0.158	0.782	0.123	0.622
$p_{10,11}$	-0.148	0.857	0.434	1.260	0.705	1.294	-0.319	1.366	0.202	0.857	0.403	0.852
$p_{11,11}$	0.079	1.256	-0.522	0.653	-0.351	0.351	-1.143	1.143	0.072	0.902	0.209	0.764

Table 5.9 95% Confidence Intervals for the Restricted OLS Estimates of the Transition Probabilities (continued)

District	Porto		Santarem		Setubal		V.Castelo		V.Real		Viseu	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	-0.066	0.368	-0.073	0.393	0.032	0.380	-0.104	0.381	-0.097	0.340	-0.092	0.368
p_{56}	0.438	0.815	0.567	1.113	0.493	1.026	0.139	0.702	0.062	1.043	0.247	0.707
p_{66}	0.122	0.431	-0.236	0.344	-0.066	0.424	0.198	0.842	-0.079	0.766	0.198	0.691
p_{67}	0.321	1.127	0.305	1.143	0.465	1.177	0.193	0.767	0.349	0.963	0.393	0.718
p_{77}	-0.220	0.572	-0.347	0.677	-0.171	0.560	-0.037	0.469	-0.100	0.586	-0.032	0.233
p_{78}	0.409	1.239	0.545	1.125	0.435	1.176	0.388	1.180	0.514	1.001	0.586	1.213
p_{88}	-0.432	0.432	-0.312	0.312	-0.349	0.434	-0.315	0.315	-0.131	0.298	-0.323	0.323
p_{89}	0.704	1.112	0.511	1.102	0.679	1.236	0.514	1.202	0.693	1.140	0.776	1.223
p_{99}	-0.249	0.249	-0.323	0.323	-0.341	0.341	-0.350	0.393	-0.209	0.283	-0.275	0.275
$p_{9,10}$	-0.342	0.836	-0.335	0.934	-0.413	0.450	-0.575	1.273	0.059	0.768	-0,146	0.484
$p_{10,10}$	0.233	0.850	-0.053	0.651	0.728	1.141	-0.218	0.511	0.065	0.439	0.560	0.853
$p_{10,11}$	0.159	0.758	0.058	0.744	-0.398	0.529	-0.111	1.299	0.373	1.123	-0.064	0.650
$p_{11,11}$	0.212	0.879	0.231	0.967	0.412	1.520	-0.407	1.186	-0.063	0.710	0.348	1.168

probability estimates, contain the corresponding time-series of the observed point estimates. This fact is not due, however, to closer estimates being obtained, but results from higher standard errors and consequently wider confidence intervals. Also, in all districts some of the transition probability estimates show large differences when compared with the observed point estimates. In particular, for the transition probabilities associated with grades 10 and 11 and in more than half of the districts, the observed point estimates do not fall within the bounds of the corresponding 95% confidence intervals.

Just as it has been observed for the whole country, the application of the basic Markov Model to the districts shows that the assumption of constant transition probabilities, for each district, and for most of the parameters , yields estimates which are not what one have hoped for, even when the restricted OLS estimation is performed. The attempt at separating the returnee students from the observed data in order to get 'corrected data matrices' will be made in the next section. As discussed in the previous chapter for the whole country, it will initially be assumed that the estimates obtained are affected by the disturbances observed in the data time-series associated with strong fluctuations in the observed point estimates of the transition probabilities, and that these disturbances are a consequence of accommodating the students transferred from the old colonies after the revolution.

5.4. The OLS Estimator for the Smoothed Data

The iterative process developed in the previous chapter, to estimate the number of returnee students allocated into the public sector school system in 1974/75 and 1975/76, was applied to all districts.

The corresponding matrices of returnee enrolment were generated and smoothed matrices of school enrolment obtained. Under the hypothesis previously established, the returnee matrices were computed assuming that, before the revolution, the educational system had a stable behaviour, with constant transition probabilities equal to the mean values of the observed point estimates during the three years prior to the revolution (see Table C.2 in Appendix C).

The iterative process applied to the whole country has shown that the restricted OLS estimates for the transition probabilities obtained using the smoothed matrix A1 (the matrix generated after the second run) are not much closer than the restricted OLS estimates obtained using the smoothed matrix A (the matrix generated after the first run). Therefore, in this regional level analysis, the process of smoothing the data matrices was stopped after the first run and consequently only one smoothed matrix of school enrolment has been generated, for each district.

Table C.3 and Table C.4, presented in Appendix C, compare the number of repeaters in grade 5 and the total number of drop-outs observed, with the corresponding values generated by the iterative process. Though smoother time-series of values are observed for the adjusted data, the same pattern and the same fluctuations noted in the observed data still remain after deducting the returnee students. FORTRAN program RETUR1 was then used for each district, to produce smoother matrices of the school enrolment. These estimated matrices of the adjusted data are presented in Appendix C, Tables C.5.1 - C.5.r. Unrestricted and restricted OLS estimators were applied to the smoothed matrices to produce the estimates of the different transition probabilities. These values are presented in Table 5.10.

Table 5.10 The OLS Estimates of the Transition Probabilities (Smoothed Data)

District	Aveiro		Beja		Braga		Braganca		C.Branco		Coimbra	
Method	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS
p_{55}	.15	.15	.16	.16	.17	.18	.12	.13	.15	.15	.16	.17
p_{56}	.67	.47	.77	.64	.35	.40	.71	.57	.79	.42	.86	.60
p_{66}	.16	.44	.10	.27	.56	.52	.17	.36	.09	.54	.01	.34
p_{67}	.44	.51	.74	.53	.39	.48	1.07	.64	.54	.46	.44	.66
p_{77}	.38	.28	.05	.30	.40	.28	-.34	.26	.25	.39	.51	.24
p_{78}	.72	.72	.51	.67	.55	.72	.56	.73	.59	.61	.58	.76
p_{88}	.14	.15	.40	.19	.31	.13	.36	.19	.34	.34	.34	.12
p_{89}	.83	.82	.81	.81	.74	.84	.80	.81	.49	.51	.82	.88
p_{99}	-.03	.00	.13	.12	.09	.00	.27	.28	.50	.49	.10	.01
$p_{9,10}$	.28	.15	.18	.16	.19	.49	.36	.45	.53	.51	-.01	.12
$p_{10,10}$	.51	.57	.61	.67	.75	.08	.16	.00	.13	.21	1.15	.89
$p_{10,11}$	.80	.43	.91	.25	1.05	.92	.25	.49	.34	.79	.36	.11
$p_{11,11}$	.40	1.00	.28	.97	.02	.01	.92	.65	.82	.35	.76	1.00
p_{5d}	-.09	.38	1.10	.20	2.43	.42	1.25	.30	1.17	.43	-.32	.23
p_{6d}	1.29	.05	-.44	.21	-4.70	.00	-2.35	.00	-1.73	.00	1.54	.00
p_{7d}	-1.54	.00	.07	.03	1.37	.00	2.06	.01	.42	.00	-1.50	.00
p_{8d}	-.35	.03	-1.45	.00	.15	.03	-1.32	.00	.16	.15	-.16	.00
p_{9d}	2.02	.85	1.54	.72	1.66	.51	.62	.27	1.64	.00	1.51	.87
$p_{10,d}$	1.59	.00	.34	.08	-.30	.00	1.63	.51	-.25	.00	1.55	.00
$p_{11,d}$	-1.08	.00	-.94	.03	4.28	.99	-.40	.35	-.26	.65	-1.42	.00

Table 5.10 The OLS Estimates of the Transition Probabilities (Smoothed Data)

(continued)

District	Evora		Faro		Guarda		Leiria		Lisboa		Portalegre	
Method	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS
p_{55}	.17	.17	.17	.20	.14	.15	.14	.18	.17	.19	.15	.22
p_{56}	1.02	.83	.76	.70	.78	.58	.67	.67	.84	.73	.57	.71
p_{66}	-.19	.06	.17	.26	.11	.37	.24	.27	.66	.22	.36	.27
p_{67}	.23	.66	.47	.74	.54	.63	.38	.73	.44	.78	.62	.73
p_{77}	.75	.19	.50	.21	.18	.07	.48	.00	.57	.21	.07	.01
p_{78}	.76	.81	.60	.79	.80	.93	.70	.23	.66	.79	.56	.99
p_{88}	.11	.05	.30	.09	.13	.00	.16	.63	.23	.09	.36	.00
p_{89}	.82	.95	.79	.91	.88	1.00	.89	.36	.93	.90	.53	.33
p_{99}	.20	.05	.12	.00	.12	.00	-.04	.52	-.04	.00	.46	.20
$p_{9,10}$	.38	.34	.25	.50	.37	.23	.21	.48	-.33	.36	.11	.00
$p_{10,10}$	.31	.38	.59	.10	.43	.60	.61	.00	1.60	.43	.94	.97
$p_{10,11}$	.71	.62	.92	.90	.30	.40	1.00	.00	-.85	.57	.42	.03
$p_{11,11}$	.40	.48	.06	.00	.85	.69	.07	.00	.23	.45	.75	1.00
p_{5d}	.31	.00	1.71	.10	.88	.26	-.39	.15	1.22	.08	1.67	.06
p_{6d}	1.81	.28	.19	.00	2.15	.00	2.09	.00	-.54	.00	-1.15	.00
p_{7d}	-2.61	.00	-1.72	.00	1.67	.00	-2.16	.77	-.05	.00	-.45	.00
p_{8d}	-1.01	.00	-.71	.00	-1.35	.77	.53	.00	-1.10	.00	.42	.67
p_{9d}	1.51	.61	2.07	.50	2.14	.00	1.42	.00	2.52	.64	.22	.80
$p_{10,d}$	-.04	.00	-3.24	.00	.42	.32	-2.05	1.00	-4.49	.00	-2.74	.00
$p_{11,d}$	1.02	.52	.97	1.00	1.18	.00	1.20	1.00	2.23	.55	2.29	.00

Table 5.10 **The OLS Estimates of the Transition Probabilities (Smoothed Data)**
(continued)

District	Porto		Santarem		Setubal		V.Castelo		V.Real		Viseu	
Method	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS	unr. OLS	res. OLS
p_{55}	.17	.19	.15	.17	.20	.21	.13	.14	.15	.16	.13	.14
p_{56}	.55	.56	.91	.65	.80	.74	.57	.77	.63	.42	.69	.46
p_{66}	.35	.36	-.05	.29	.11	.20	.31	.05	.20	.51	.14	.47
p_{67}	.55	.64	.77	.71	.65	.80	.36	.44	.17	.49	.43	.53
p_{77}	.35	.25	.08	.17	.36	.20	.38	.24	.86	.43	.29	.13
p_{78}	.74	.75	.61	.83	.76	.80	1.07	.76	.40	.57	.56	.87
p_{88}	.07	.08	.26	.00	.08	.05	-.49	.00	.53	.32	.38	.02
p_{89}	.93	.92	.77	.85	1.02	.95	.50	.40	.62	.68	1.18	.98
p_{99}	.05	.00	.15	.00	-.09	.00	.48	.59	.34	.29	-.25	.00
$p_{9,10}$	.06	.46	.05	.00	-.03	.38	.23	.41	.41	.22	.11	.08
$p_{10,10}$	.94	.12	.99	.94	1.12	.20	.57	.00	.23	.68	.89	.91
$p_{10,11}$	.99	.88	1.00	.06	.96	.80	1.25	1.00	-.91	.32	.45	.09
$p_{11,11}$	.07	.08	-.02	.94	.07	.16	-.01	.00	2.43	1.00	.68	1.00
p_{5d}	1.02	.25	1.92	.18	-1.59	.04	-3.20	.09	1.16	.42	1.04	.40
p_{6d}	-.51	.00	-3.44	.00	6.66	.00	9.27	.51	.19	.00	-.64	.00
p_{7d}	-.86	.00	.38	.00	-4.39	.00	-7.84	.00	-2.21	.00	.91	.00
p_{8d}	.28	.00	.46	.15	-.15	.00	-2.34	.60	-.51	.00	-1.89	.00
p_{9d}	.93	.54	2.29	1.00	.71	.62	1.63	.00	1.31	.49	1.17	.92
$p_{10,d}$	-1.91	.00	3.11	.00	-4.62	.00	-2.35	.00	-1.22	.00	-4.83	.00
$p_{11,d}$	2.19	.92	-2.94	.05	3.50	.83	4.44	1.00	2.92	.00	4.97	.00

The FORTRAN program RESIDD was also used to give the statistics for the restricted OLS estimates. Table 5.11 shows the t-values for the unrestricted and restricted OLS estimates of the transition probabilities and Table 5.12 and Table 5.13 present the corresponding 95% confidence intervals.

Even after removing the estimated number of returnee students from the original data, the regional analysis shows that the unrestricted OLS estimates still do not satisfy the non-negativity and the row-sum conditions. A comparison between the unrestricted OLS and the restricted OLS estimators computed reveal that, although the results differ from one district to another, there is no district for which one process gives all estimates closer to the assumed constant point estimates. Most of the equations of the model still show that only one of the independent variables has statistical significance. Though this significance is higher in the case of the restricted OLS (see Table 5.11), lower t-values still occur, in many districts, for estimates corresponding to the last grade or to the first grade of the secondary levels of education.

As observed in the previous section for the analysis using the original data, the 95% confidence intervals for the restricted OLS estimates presented in Table 5.13 are, globally, smaller than those for the unrestricted OLS estimates presented in Table 5.12. These tables show that in only three districts (Faro, Leiria and V. do Castelo) the 95% confidence intervals for the set of the unrestricted and restricted OLS transition probability estimates contain the corresponding constant point estimates. For the unrestricted OLS estimator, two other districts (Braga and V.Real) have 95% confidence intervals containing the constant point estimates, and for the

Table 5.11 T-values for the OLS Estimates of the Transition Probabilities for the Smoothed Data

District	Aveiro		Beja		Braga		Braganca		C.Branco		Coimbra		Evora		Faro		Guarda	
Method	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS
p_{55}	8.82	1.31	10.03	1.55	11.33	2.20	8.43	1.76	10.00	1.54	8.89	1.81	8.89	0.73	9.44	1.94	9.33	1.95
p_{56}	5.11	5.84	10.69	7.62	1.74	3.52	7.55	3.35	7.38	3.75	3.66	4.28	2.47	3.02	3.19	6.03	4.46	6.90
p_{66}	0.94	6.17	1.12	3.81	1.89	5.85	1.41	4.14	0.69	5.63	0.03	2.84	0.39	0.17	0.61	2.65	0.51	6.17
p_{67}	2.30	2.45	4.48	2.45	2.03	2.26	7.70	3.80	3.97	2.74	1.78	3.61	1.41	3.19	2.06	3.03	3.18	3.28
p_{77}	1.29	1.55	0.21	1.77	1.24	2.12	1.86	1.56	1.24	2.69	1.70	1.70	3.41	0.81	1.89	1.07	0.66	0.58
p_{78}	4.65	4.59	3.45	2.85	3.40	3.94	5.28	4.17	4.61	2.88	3.49	4.27	3.45	3.55	4.03	3.41	7.62	7.88
p_{88}	0.72	0.82	2.05	0.64	1.45	0.68	2.75	1.27	2.25	2.71	1.69	0.65	0.42	0.21	1.63	0.35	1.07	0.00
p_{89}	9.54	10.51	6.64	7.36	5.69	7.08	5.52	6.38	3.45	3.05	9.01	10.23	3.19	5.60	5.77	5.41	10.35	13.61
p_{99}	0.27	0.00	0.94	0.82	0.54	0.00	1.96	1.33	3.13	2.68	0.95	0.09	0.75	0.21	0.75	0.00	1.36	0.00
$p_{9,10}$	2.30	0.61	4.39	1.54	1.11	2.38	2.87	2.95	3.71	2.85	0.12	0.81	1.48	1.22	1.23	1.44	4.02	1.45
$p_{10,10}$	2.10	5.11	5.50	19.14	2.43	0.70	1.61	0.00	0.51	1.71	7.77	13.22	0.62	2.53	2.01	0.50	2.69	6.68
$p_{10,11}$	2.31	1.69	3.97	1.06	3.39	4.84	0.53	2.84	2.28	5.56	1.30	0.83	1.47	2.44	3.00	2.88	1.94	3.74
$p_{11,11}$	1.17	2.91	1.26	2.98	0.06	0.04	2.70	3.38	5.16	1.81	2.55	5.17	0.83	1.53	0.18	0.00	5.31	5.28

Table 5.11 T-values for the OLS Estimates of the Transition Probabilities for the Smoothed Data (continued)

District	Leiria		Lisboa		Portalegre		Porto		Santarem		Setubal		V.Castelo		V.Real		Viseu	
Method	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS	unrest OLS	rest OLS
p_{55}	9.33	2.12	14.17	2.64	10.01	1.17	10.00	2.02	10.00	0.88	11.76	2.73	9.29	1.34	8.82	1.82	8.13	1.38
p_{56}	5.32	7.98	6.67	14.04	2.45	3.23	4.10	8.62	4.97	6.19	2.50	6.61	2.44	4.81	3.41	4.00	5.52	4.45
p_{66}	2.48	3.86	4.43	4.49	1.30	1.42	1.99	7.50	0.22	2.87	0.29	2.13	0.99	0.28	0.78	6.40	0.82	4.52
p_{67}	1.85	2.98	1.59	4.02	6.26	3.17	2.14	4.60	2.75	3.64	2.25	6.35	3.36	4.63	0.90	4.45	4.17	6.56
p_{77}	1.53	0.00	1.88	1.43	4.03	0.05	1.06	2.25	0.23	0.76	1.15	1.69	1.80	3.43	3.23	4.03	1.53	2.13
p_{78}	5.07	0.88	5.55	3.76	3.24	2.54	3.54	3.07	4.77	5.34	4.20	4.26	4.44	3.36	2.45	2.97	2.32	4.96
p_{88}	0.92	2.39	1.04	0.36	1.71	0.00	0.26	0.32	1.57	0.00	0.34	0.25	1.44	0.00	2.36	2.71	1.28	0.12
p_{89}	7.06	1.85	5.74	6.91	5.52	0.60	6.64	7.30	7.40	8.09	5.13	6.79	1.89	2.12	3.48	5.76	9.75	1.04
p_{99}	0.24	2.39	0.21	0.00	4.22	0.23	0.31	0.00	1.22	0.00	0.41	0.00	1.57	2.92	3.02	1.93	1.89	0.00
$p_{9,10}$	1.12	1.81	1.88	1.67	1.17	0.00	0.58	1.59	0.29	.00	0.71	1.82	1.24	0.78	3.87	1.20	1.49	0.54
$p_{10,10}$	1.66	0.00	5.63	3.64	5.80	7.03	4.92	0.98	2.93	8.45	3.01	2.25	1.14	0.00	1.06	6.53	5.78	15.90
$p_{10,11}$	1.95	0.00	2.21	3.26	1.64	0.17	3.64	5.00	3.67	0.28	2.07	3.92	3.28	2.79	1.12	1.25	1.53	0.54
$p_{11,11}$	0.14	0.00	0.23	1.85	2.74	4.39	0.24	0.33	0.02	4.07	0.14	0.63	0.03	0.00	2.45	4.47	2.24	5.15

Table 5.12 95% Confidence Intervals for the Unrestricted OLS Estimates of the Transition Probabilities for the Smoothed Data

District	Aveiro		Beja		Braga		Braganca		C.Branco		Coimbra	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	0.114	0.187	0.120	0.193	0.132	0.200	0.085	0.148	0.113	0.181	0.123	0.204
p_{56}	0.376	0.969	0.607	0.933	-0.106	0.805	0.192	0.923	0.547	1.029	0.331	1.394
p_{66}	-0.232	0.559	-0.102	0.298	-0.107	1.232	-0.322	0.446	-0.201	0.386	-0.637	0.664
p_{67}	0.004	0.870	0.347	1.093	0.046	0.823	0.761	0.392	0.230	0.845	-0.116	1.001
p_{77}	-0.286	1.046	-0.540	0.548	-0.333	1.127	-0.075	0.074	-0.192	0.721	-0.173	1.195
p_{78}	0.367	1.071	0.171	0.843	0.184	0.916	0.326	0.803	0.302	0.883	0.206	0.957
p_{88}	-0.296	0.581	-0.044	0.839	-0.170	0.798	0.061	0.653	-0.001	0.664	-0.120	0.791
p_{89}	0.636	1.029	0.534	1.087	0.451	1.037	0.471	1.128	0.168	0.810	0.610	1.020
p_{99}	-0.284	0.221	-0.183	0.443	-0.291	0.464	-0.040	0.586	0.139	0.864	-0.141	0.336
$p_{9,10}$	0.008	0.560	0.091	0.278	-0.200	0.575	0.148	0.573	0.208	0.854	-0.198	0.173
$p_{10,10}$	-0.040	1.061	0.364	0.865	0.047	1.444	0.346	0.670	-0.440	0.713	0.815	1.485
$p_{10,11}$	-0.063	1.503	0.389	1.424	0.345	1.749	-0.437	0.932	0.001	0.676	-0.267	0.988
$p_{11,11}$	-0.375	1.167	-0.222	0.788	-0.740	0.787	0.148	1.689	0.461	1.180	0.085	1.431

Table 5.12 95% Confidence Intervals for the Unrestricted OLS Estimates of the Transition Probabilities for the Smoothed Data (continued)

District	Evora		Faro		Guarda		Leiria		Lisboa		Portalegre	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	0.122	0.208	0.134	0.212	0.105	0.173	0.107	0.176	0.144	0.196	0.117	0.186
p_{56}	0.094	1.983	0.224	1.301	0.386	1.177	0.373	0.945	0.557	1.129	0.046	0.991
p_{66}	-1.293	0.921	-0.460	0.797	-0.384	0.596	-0.123	0.604	0.171	0.404	-0.263	0.991
p_{67}	-0.134	0.603	-0.046	0.984	0.155	0.926	-0.082	0.847	-0.186	1.066	0.396	0.844
p_{77}	0.256	1.250	-0.096	1.102	-0.434	0.799	-0.229	1.187	-0.114	1.257	-0.285	0.423
p_{78}	0.266	1.262	0.262	0.937	0.558	1.034	0.387	1.014	0.257	1.070	0.171	0.952
p_{88}	-0.487	0.715	-0.116	0.718	-0.146	0.403	-0.233	0.556	-0.276	0.728	-0.113	0.843
p_{89}	0.235	1.397	0.483	1.101	0.690	1.073	0.604	1.175	0.561	1.295	0.312	0.744
p_{99}	-0.402	0.795	-0.239	0.481	-0.075	0.323	-0.382	0.302	-0.472	0.389	0.215	0.710
$p_{9,10}$	-0.204	0.958	-0.112	0.621	0.162	0.577	-0.216	0.636	-0.730	0.068	-0.103	0.321
$p_{10,10}$	-0.829	1.453	-0.070	1.259	0.071	0.795	-0.221	1.442	0.953	2.238	0.575	1.309
$p_{10,11}$	-0.387	1.798	0.224	1.614	-0.047	0.654	-0.163	2.155	-0.201	1.718	-0.159	1.000
$p_{11,11}$	-0.699	1.491	-0.702	0.827	0.490	1.214	-1.064	1.194	-0.658	1.115	0.131	1.373

Table 5.12 95% Confidence Intervals for the Unrestricted OLS Estimates of the Transition Probabilities for the Smoothed Data (continued)

District	Porto		Santarem		Setubal		V.Castelo		V.Real		Viseu	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	0.131	0.209	0.120	0.185	0.163	0.236	0.094	0.158	0.112	0.191	0.098	0.170
p_{56}	0.247	0.855	0.495	1.324	0.083	1.529	0.037	1.096	0.212	1.050	0.406	0.971
p_{66}	-0.047	0.751	-0.555	0.452	-0.760	0.977	-0.402	1.013	-0.376	0.781	-0.247	0.529
p_{67}	-0.030	1.131	0.136	1.401	-0.007	1.299	0.112	0.598	-0.254	0.603	0.199	0.666
p_{77}	-0.400	1.092	-0.711	0.860	-0.346	1.075	-0.098	0.855	0.263	1.465	-0.141	0.719
p_{78}	0.266	1.213	0.319	0.897	0.350	1.167	0.528	1.618	0.030	0.767	0.011	1.100
p_{88}	-0.544	0.686	-0.114	0.637	-0.444	0.609	-1.260	0.285	0.024	1.041	-0.290	1.053
p_{89}	0.613	1.245	0.532	1.000	0.571	1.471	-0.096	1.105	0.215	1.019	0.906	1.454
p_{99}	-0.415	0.317	-0.134	0.424	-0.585	0.412	-0.214	1.171	-0.120	0.809	-0.546	0.053
$p_{9,10}$	-0.174	0.290	-0.344	0.445	-0.418	0.364	-0.195	0.647	0.174	0.652	0.058	0.277
$p_{10,10}$	0.509	1.375	0.220	1.750	0.281	1.966	-0.570	1.704	-0.257	0.721	0.544	1.242
$p_{10,11}$	0.376	1.607	0.385	1.615	-0.093	2.005	0.384	2.107	-2.740	0.926	-0.214	1.117
$p_{11,11}$	-0.582	0.714	-0.749	0.706	-1.056	1.198	-0.825	0.810	0.188	4.667	-0.108	1.363

Table 5.13 95% Confidence Intervals for the Restricted OLS Estimates of the Transition Probabilities for the Smoothed Data

District	Aveiro		Bega		Braga		Braganca		C.Branco		Coimbra	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	-0.109	0.409	-0.085	0.385	-0.087	0.447	-0.038	0.298	-0.071	0.371	-0.044	0.383
p_{56}	0.284	0.656	0.446	0.834	0.019	0.781	0.321	0.819	0.162	0.678	0.276	0.924
p_{66}	0.275	0.605	0.107	0.433	0.222	0.818	0.160	0.560	0.318	0.762	0.064	0.617
p_{67}	0.031	0.989	0.031	1.029	-0.233	1.193	0.251	1.029	0.072	0.848	0.237	1.083
p_{77}	-0.138	0.698	-0.089	0.689	-0.164	0.723	-0.124	0.644	0.056	0.724	-0.085	0.565
p_{78}	0.358	1.082	0.128	1.212	0.107	1.333	0.326	1.134	0.121	1.099	0.350	1.170
p_{88}	-0.274	0.574	-0.494	0.874	-0.508	0.768	-0.157	0.537	-0.094	0.774	-0.306	0.546
p_{89}	0.639	1.000	0.555	1.065	0.442	1.238	0.516	1.104	0.125	0.895	0.682	1.078
p_{99}	-0.175	0.175	-0.216	0.456	-0.401	0.401	-0.278	0.695	0.069	0.911	-0.236	0.256
$p_{9,10}$	-0.415	0.715	-0.081	0.401	-0.202	1.182	0.098	0.801	0.097	0.923	-0.221	0.461
$p_{10,10}$	0.313	0.827	0 588	0.752	-0.303	0.463	-0.143	0.143	-0.073	0.493	0.735	1.045
$p_{10,11}$	-0.157	1.017	-0.292	0.792	0.282	1.558	0.090	0.890	0.462	1.118	-0.194	0.414
$p_{11,11}$	0.207	1.793	0.219	1.721	-0.738	0.758	0.207	1.093	-0.096	0.796	0.603	1.397

Table 5.13 95% Confidence Intervals for the Restricted OLS Estimates of the Transition Probabilities for the Smoothed Data (continued)

District	Evora		Faro		Guarda		Leiria		Lisboa		Portalegre	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	-0.361	0.701	-0.037	0.429	-0.024	0.324	-0.013	0.373	0.024	0.352	-0.206	0.646
p_{56}	0.197	1.463	0.435	0.970	0.385	0.775	0.476	0.864	0.611	0.849	0.203	1.217
p_{66}	-0.756	0.876	0.036	0.490	0.171	0.569	0.108	0.432	0.105	0.330	-0.168	0.709
p_{67}	0.183	1.137	0.174	1.299	0.186	1.074	0.164	1.296	0.336	1.229	0.198	1.262
p_{77}	-0.352	0.731	-0.240	0.664	-0.209	0.350	-0.408	0.408	-0.126	0.551	-0.422	0.442
p_{78}	0.284	1.336	0.250	1.323	0.657	1.203	-0.374	0.834	0.285	1.290	0.091	1.889
p_{88}	-0.506	0.606	-0.461	0.647	-0.285	0.285	0.021	1.239	-0.486	0.667	-0.640	0.640
p_{89}	0.559	1.341	0.520	1.295	0.831	1.169	-0.089	0.809	0.601	1.218	-0.944	1.604
p_{99}	-0.507	0.607	-0.447	0.447	-0.209	0.209	0.017	1.023	-0.389	0.389	-1.802	2.202
$p_{9,10}$	-0.302	0.982	-0.302	1.296	-0.136	0.596	-0.134	1.093	-0.136	0.861	-0.400	0.400
$p_{10,10}$	0.033	0.727	-0.363	0.555	0.393	0.807	-0.272	0.272	0.153	0.697	0.666	1.274
$p_{10,11}$	0.034	1.206	0.181	1.627	0.152	0.648	-1.523	1.523	0.172	0.978	-0.380	0.440
$p_{11,11}$	-0.377	1.337	-1.005	1.005	0.389	0.991	-2.140	2.140	-0.112	1.007	0.474	1.526

Table 5.13 95% Confidence Intervals for the Restricted OLS Estimates of the Transition Probabilities for the Smoothed Data (continued)

District	Porto		Santarem		Setubal		V. Castelo		V. Real		Viseu	
	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.	L.B.	U.B.
p_{55}	-0.026	0.397	-0.087	0.427	0.038	0.384	-0.970	0.377	-0.040	0.360	-0.089	0.369
p_{56}	0.411	0.711	0.408	0.892	0.482	0.999	0.400	1.140	0.178	0.622	0.222	0.698
p_{66}	0.245	0.466	0.056	0.524	-0.022	0.414	-0.356	0.456	0.326	0.694	0.230	0.710
p_{67}	0.324	0.965	0.261	1.159	0.514	1.094	0.222	0.658	0.235	0.745	0.344	0.716
p_{77}	-0.012	0.503	-0.345	0.685	-0.072	0.474	0.078	0.402	0.184	0.676	-0.011	0.271
p_{78}	0.229	1.280	0.471	1.189	0.366	1.231	0.237	1.283	0.128	1.012	0.465	1.275
p_{88}	-0.488	0.650	-0.373	0.373	-0.415	0.506	-0.382	0.382	0.048	0.592	-0.353	0.393
p_{89}	0.628	1.210	0.607	1.693	0.631	1.278	-0.035	0.835	0.408	0.952	0.763	1.197
p_{99}	-0.384	0.384	-0.286	0.286	-0.417	0.417	0.125	1.055	-0.057	0.637	-0.269	0.269
$p_{9,10}$	-0.214	1.127	-0.493	0.493	-0.106	0.860	-0.796	1.616	-0.201	0.641	-0.261	0.421
$p_{10,10}$	-0.159	0.405	0.684	1.196	-0.008	0.403	-0.414	0.414	0.440	0.920	0.778	1.042
$p_{10,11}$	0.469	1.280	-0.439	0.559	0.333	1.272	0.175	1.825	-0.271	0.911	-0.298	0.478
$p_{11,11}$	-0.470	0.639	0.407	1.473	-0.420	0.746	-1.276	1.276	0.484	1.516	0.553	1.447

restricted OLS estimator, three other districts (Braganca, Evora and Setubal) give a similar result.

Tables C.6.a - C.6.r, presented in Appendix C, summarise the results obtained in this chapter, showing the different estimates computed for the transition probabilities. It is clear that even for Faro, Leiria and V. do Castelo, some of the estimates obtained are unreal and so not acceptable, either using one method, or using the other, or using both methods.

A graphical representation of the restricted OLS estimates of the transition probabilities, using the observed data and the adjusted data, is presented in Fig 5.1 to Fig 5.18. The results differ from one district to another. Some districts (Coimbra, Faro, Lisboa, Viseu) present similar graph patterns and others (Guarda, Leiria, Portalegre and V. do Castelo) present completely different graphs. Also, for this last set of districts, none of the restricted OLS procedures give acceptable estimates for the transition probabilities. For the remaining group of districts, the process of smoothing the data has generated, in some districts, closer values for the transition probability estimates in the secondary level; for other districts, the transition probability estimates are more reasonable only in the unified general course.

In conclusion, the attempt at smoothing the data matrices of school enrolment has led to either similar or more reasonable values for the restricted OLS transition probability estimates in only five districts (Evora, Faro, Lisboa, Porto and Setubal).

Fig.5.1-5.18 -The Restricted OLS Estimates for the Observed Data and for the Smoothed Data by District

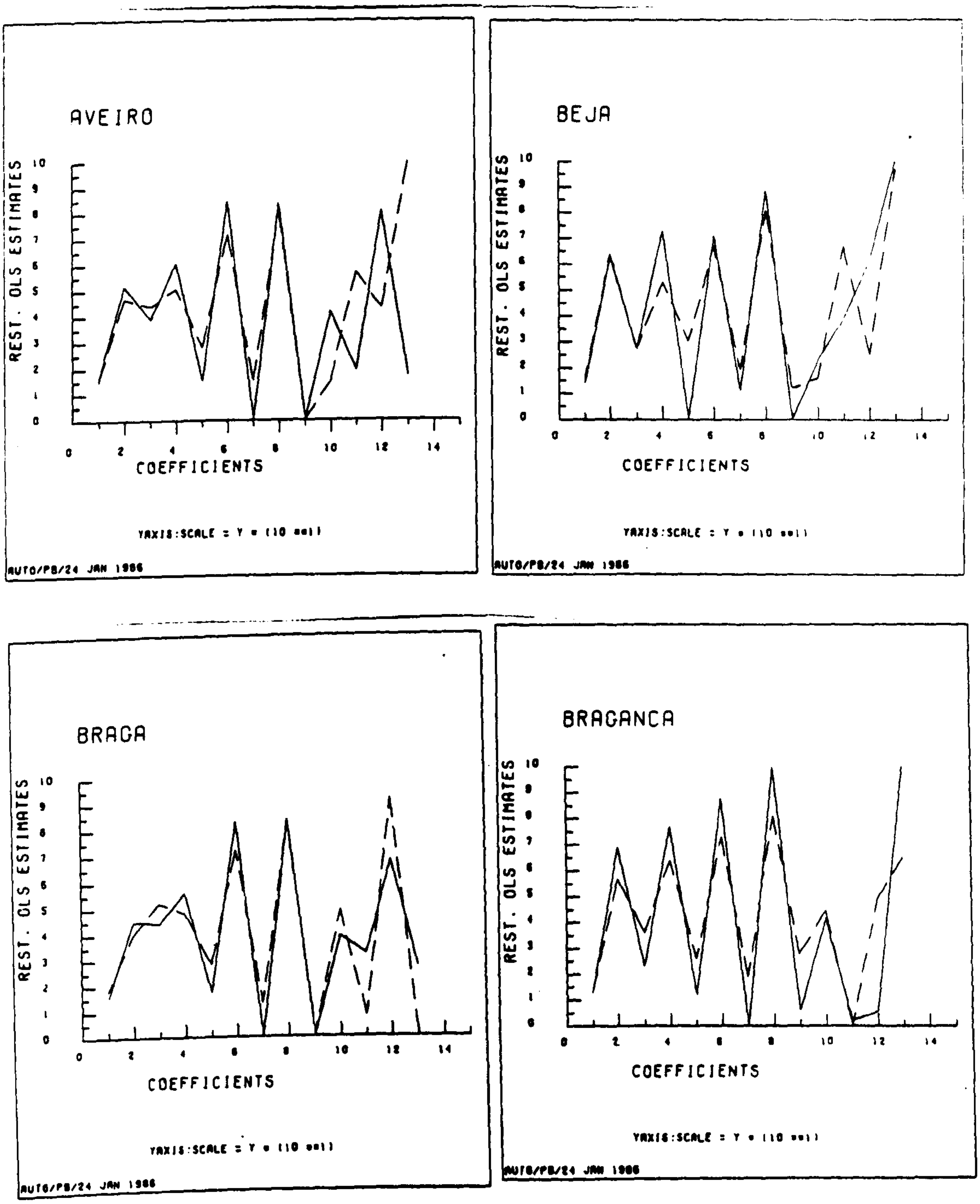


Fig.5.1-5.18 -The Restricted OLS Estimates for the Observed Data and for the Smoothed Data by District (Continued)

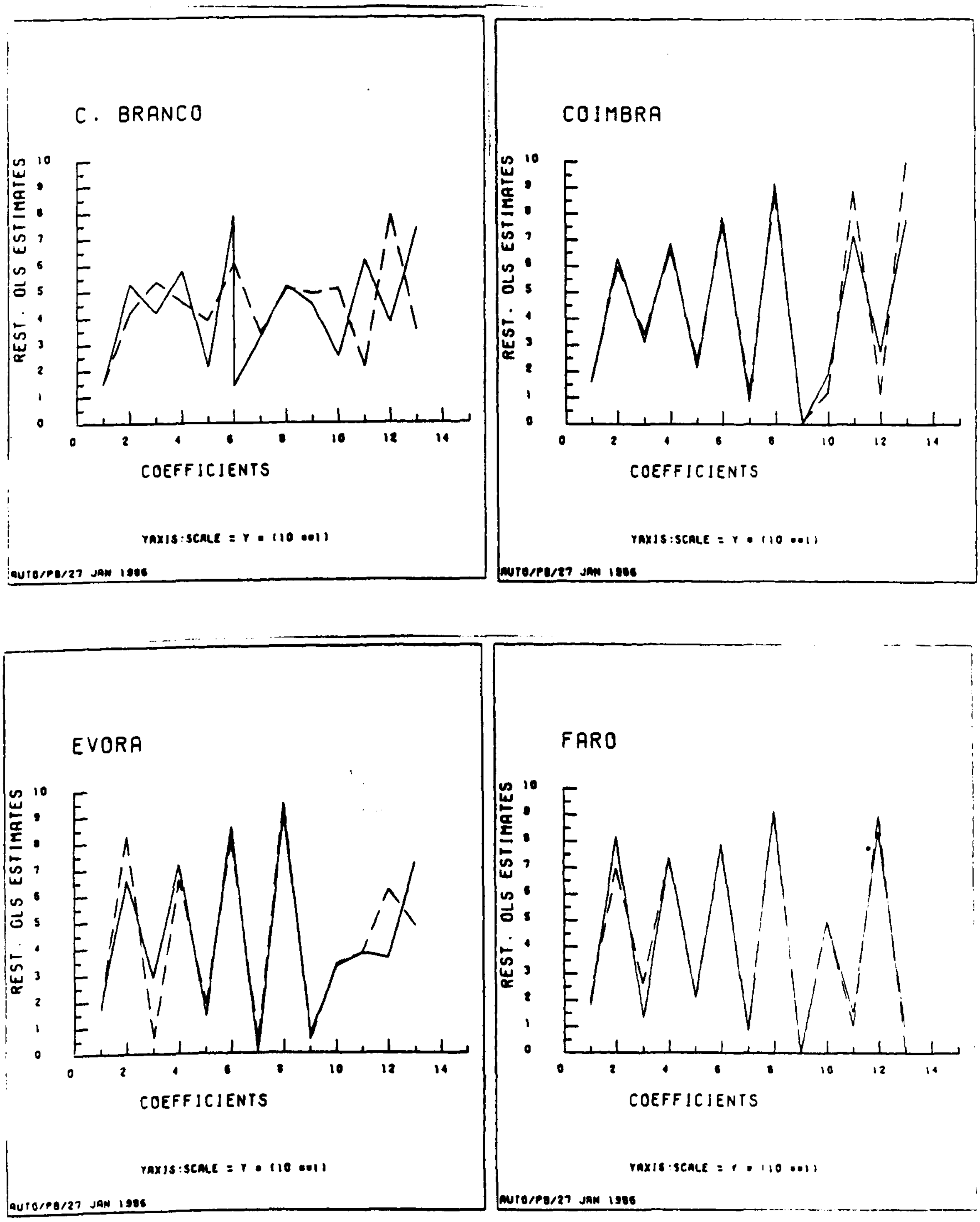


Fig.5.1-5.18 -The Restricted OLS Estimates for the Observed Data and for the Smoothed Data by District (Continued)

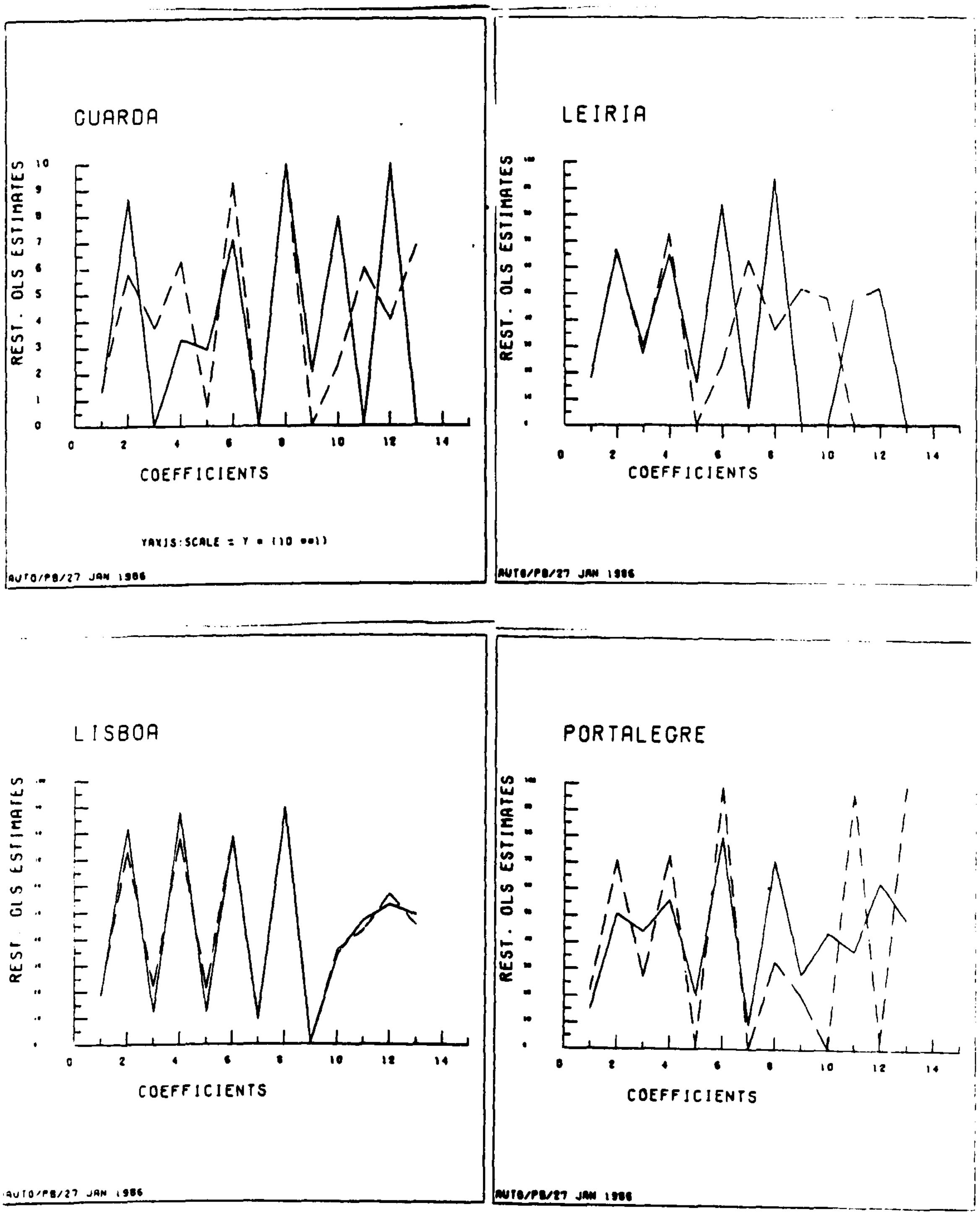


Fig.5.1-5.18 -The Restricted OLS Estimates for the Observed Data and for the Smoothed Data by District (Continued)

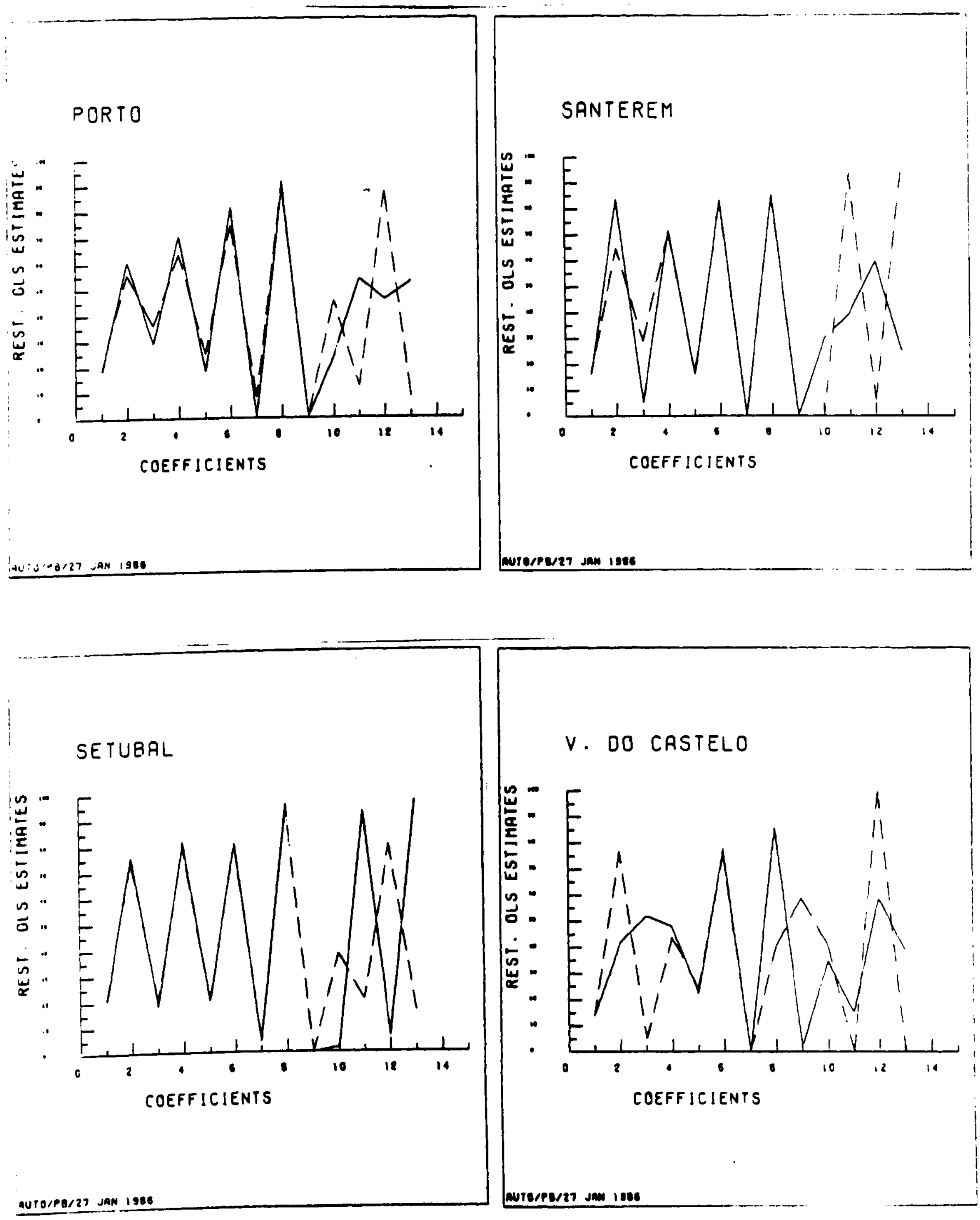
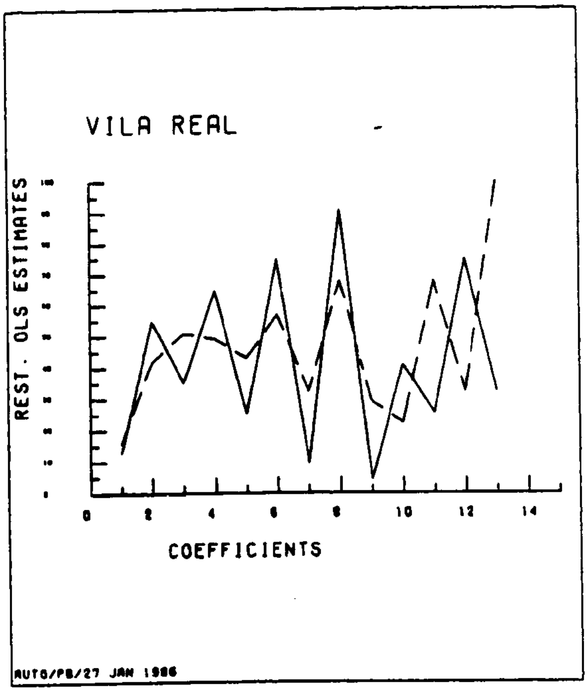
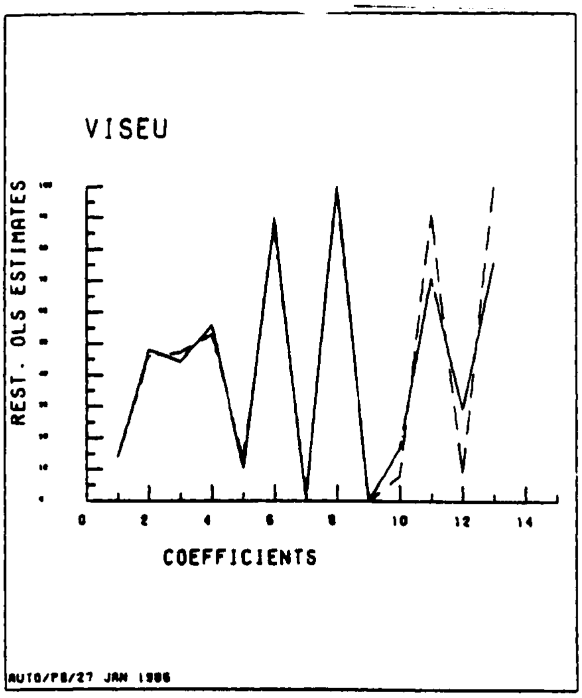


Fig.5.1-5.18 -The Restricted OLS Estimates for the Observed Data and for the Smoothed Data by District (Continued)





——— Observed Data

— — — Smoothed Data

Thus a brief note on the results can be made: the application of the basic Markov model described in Chapter 3 to the Portuguese educational system and developed in the previous chapter has proved to be inappropriate, as the transition probability estimates are biased and non-efficient. The attempt made in this chapter at applying the same Markov model at the district level had the aim of analysing whether or not a district-level approach would give better estimates of the different probabilities. Furthermore, a better understanding of the differences in the performance of the enrolment and the corresponding transition probabilities could result. However, it has been seen that the non-stationarity of the transition probabilities is apparent in all districts and consequently the OLS estimation procedures, when applied to the data, also give biased estimates for the transition probabilities. Even the five districts where the restricted OLS procedure gave overall better estimates, present values far from the observed point estimates in several cases.

Thus it is very difficult to identify districts with homogeneous behaviour in terms of the educational variables: enrolment and transition probabilities, as the estimates of these probabilities do not represent the true performance. Therefore, even in a more disaggregated analysis, the assumption of constant transition probabilities is limited and unreliable when trying to study the behaviour of the school enrolment of the Portuguese educational system, through the corresponding transition probabilities.

Chapter 5

Footnotes

1. One way of identifying these migratory movements is by comparing the students' place of study every year with the students' place of study the year before. This information is rarely available and Portugal is no exception.

2. Note that the number of students in private schools decreased during the 1970s.

3. Note that these districts also have lower drop-out rates.

Chapter 6

TIME-VARYING MARKOV MODELS

6.1. Introduction

The assumption of parameter stability of the Markov model clearly may be violated when applying this model to the educational system. The behaviour of the school enrolment may change, altering the causal structure underlying the process. Forecasts, even short-term ones, have turned out to be far from what actually happened in reality, suggesting that insufficient information concerning the underlying factors which affect students behaviour and change over time is input into the model. The results obtained in Chapter 4 and in Chapter 5 reinforce this statement. They have shown that it is not reasonable to assume that the transition probabilities are invariant over the sample period employed. It seems, therefore, sensible to believe that exogenous factors may have a significant influence on the probability of a student being promoted, repeating or leaving the school system.

It is the purpose of this study to identify some of the underlying factors by examining the changes of some socio-economic and school related features together with the changes in the transition probabilities. This chapter traces, therefore, the theoretical threads of a more flexible Markov model, in which the parameters vary over the sample period. A causal structure is then incorporated in the model, via a regression formulation, allowing parameters to become functions of exogenous variables. Ordinary least squares estimation processes[1] are also described to derive the estimates of the unknown transition probabilities.

It must be noted, however, that when applying a time-varying Markov model to the educational system, the time-patterns of the transition probabilities become more important than their absolute values. Transition probability estimates whose time-patterns fit the time-patterns of the corresponding observed point estimates may give useful guidelines to assist educational planners and decision makers in the preparation of the educational plans. The marginal effects of each causal variable upon the transition probabilities may then be useful indicators for policy purposes.

6.2. The Extended Model

The general equations of the first-order time varying Markov model, applied to the educational system can be written using similar expressions to those presented in Chapter 3 for the basic Markov model [see equations (3.2)]

$$\underline{n}_a'(t) = \underline{n}_a'(t-1)\ \underline{P}^t_{a,t-1} + \underline{n}_o^{*\prime}(t) \qquad (6.1)$$

with

$$\sum_j pij(t) = 1 \quad \text{and} \quad p_{ij}(t) < 1$$

$$i,j=1, \ldots ,s+1$$
$$t=1, \ldots , k$$

where s+1 is the number of states of the system, $\underline{n}_a'(t)$, $\underline{n}_o^*(t)$ are [1 x (s+1)] row vectors and $\underline{P}^t_{a,t-1}$ is the augmented square matrix [(s+1) x (s+1)] of transition probabilities $p_{ij}(t)$.

The main difference between the two sets of equations, equations (3.2) and equations (6.1), is that the transition probabilities are now variable over time and so differing k transition matrices $\underline{P}^{t}_{a,t-1}$ must be incorporated into the model. Furthermore, the transition probabilities are themselves functions of a set of explanatory variables $x_1(t), x_2(t), \ldots, x_m(t)$, and these functions are determined by new parameters a_{ij} and δ_{ij}, by:

$$p_{ij}(t) = a_{ij} + \sum_{h=1}^{m} \delta_{ijh}\, x_h(t) + v_{ij}(t) \qquad (6.2)$$

for all i, j, h

where a_{ij} is the constant term of the regression equation and $v_{ij}(t)$ is a disturbance term. It must be noted that in the basic model described by equations (3.2), there are $(s+1)^2$ unknown transition probabilities and $(s+1)k$ equations; with this new formulation $(s+1)^2k$ unknown transition probabilities exist and $(s+1)^2k$ new equations of type (6.2) are added to the previous set of $(s+1)k$ equations of type (6.1).

Following the same structure presented in Chapter 3 for the basic Markov model, the extended model can be written in the form:

$$\underline{n}^* = \underline{N}\,\underline{p} + \underline{u} \qquad (6.3)$$

$$\underline{p} = \underline{X}\,\underline{\delta} + \underline{v} \qquad (6.4)$$

where $\underline{n}^*$ is the $[(s+1)k \times 1]$ vector of the number of students enrolled by grade, $\underline{u}$ and $\underline{v}$ are the $[(s+1)k \times 1]$ and $[(s+1)^2k \times 1]$ vectors of disturbance terms, $\underline{N}$ is the $[(s+1)k \times s(s+1)k]$ block diagonal matrix of the number of students enrolled by grade in the previous year. Matrix $\underline{N}$ has, in the case of the extended model, the form

$$\underline{N} = \begin{bmatrix} \underline{N}_1 & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{N}_2 & & \underline{0} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{0} & \underline{0} & \cdots & \underline{N}_{s+1} \end{bmatrix}$$

with $\underline{N}_1 = \underline{N}_2 = \cdots = \underline{N}_{s+1}$, where each $\underline{N}_i$ is a (k x sk) matrix with the form

$$\underline{N}_i = \begin{bmatrix} n_1(0) & \cdots & 0 & n_2(0) & \cdots & 0 & & n_s(0) & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \cdots & \vdots & \ddots & \vdots \\ 0 & \cdots & n_1(k-1) & 0 & \cdots & n_2(k-1) & & 0 & \cdots & n_s(k-1) \end{bmatrix}$$

The [s(s+1)k x 1] vector $\underline{p}$ of variable transition probabilities $p_{ij}(t)$ is aggregated in the form

$$\underline{p} = \begin{bmatrix} \underline{p}^1 \\ \vdots \\ \underline{p}_{s+1} \end{bmatrix} \text{ with } \underline{p}_j = \begin{bmatrix} \underline{p}_{1j} \\ \vdots \\ \underline{p}_{sj} \end{bmatrix} \text{ and } \underline{p}_{ij} = \begin{bmatrix} p_{ij}(1) \\ \vdots \\ p_{ij}(k) \end{bmatrix}.$$

where $p_{ij}(t)$ is the probability that a student in grade i at time t-1 (school year t-1/t) achieves grade j at time t (school year t/t+1). Matrix $\underline{X}$ is a $[s(s+1)k \times (s+1)^2 m]$ block diagonal matrix of the m explanatory variables,

$$\underline{X} = \begin{bmatrix} \underline{X}_1 & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{X}_2 & \cdots & \underline{0} \\ \vdots & \vdots & \ddots & \vdots \\ \underline{0} & \underline{0} & \cdots & \underline{X}_{s(s+1)} \end{bmatrix}$$

in which $\underline{X}_1 = \underline{X}_2 = \cdots = \underline{X}_{s(s+1)}$ and each $\underline{X}_i$ is a $[k \times (m+1)]$ matrix of the form

$$\underline{X}_i = \begin{bmatrix} 1 & x_1(1) & \cdots & x_m(1) \\ \cdot & \cdots & \cdots & \cdots \\ 1 & x_1(k) & \cdots & x_m(k) \end{bmatrix}$$

This means that it is assumed that the same set of explanatory variables is used to explain the changes over time of each of the transition probabilities.

Finally, the vector $\underline{\delta}$ of the new parameters has size $[(s+1)^2 m \times 1]$ and the following form

$$\underline{\delta} = \begin{bmatrix} \underline{\delta}_1 \\ \\ \cdot \\ \cdot \\ \underline{\delta}_{s+1} \end{bmatrix} \quad \text{with} \quad \underline{\delta}_j = \begin{bmatrix} \underline{\delta}_{1j} \\ \cdot \\ \cdot \\ \cdot \\ \underline{\delta}_{s+1,j} \end{bmatrix} \quad \text{and} \quad \underline{\delta}_{ij} = \begin{bmatrix} a_{ij} \\ \delta_{ij}(1) \\ \cdot \\ \cdot \\ \cdot \\ \delta_{ij}(m) \end{bmatrix}$$

The stochastic assumptions for the disturbance terms are: (i) the expected value is zero and the variance is constant for all observations, (ii) the disturbances corresponding to different observations have zero correlation and (iii) the disturbance term is normally distributed. These assumptions can be summarised for this extended model by

$$E(\underline{u}) = \underline{0} \qquad E(\underline{v}) = \underline{0} \qquad E(\underline{u}\,\underline{v}') = \underline{0}$$

$$E(\underline{u}\,\underline{u}') = \underline{\Sigma} \qquad E(\underline{v}\,\underline{v}') = \underline{\Omega}$$

where $\underline{\Sigma}$ is a [(s+1)k x (s+1)k] block diagonal covariance matrix and $\underline{\Omega}$ is a [s(s+1)k x s(s+1)k] block diagonal covariance matrix.

Substituting equation (6.4) into equation (6.3) gives:

$$\underline{n}^* = \underline{N}\,\underline{X}\,\underline{\delta} + \underline{w} \tag{6.5}$$

where

$$\underline{w} = \underline{N}\,\underline{v} + \underline{u} \tag{6.6}$$

Equation (6.5) is then the abreviated equation of the extended Markov model, after incorporating the explanatory variables, with a disturbance term $\underline{w}$ with the following properties

(i) $E(\underline{w}) = E(\underline{N}\,\underline{v} + \underline{u}) = \underline{N}\,E(\underline{v}) + E(\underline{u}) = \underline{0}$

(ii)
$$\begin{aligned} E(\underline{w}\,\underline{w}') &= E[(\underline{N}\,\underline{v} + \underline{u})\,(\underline{N}\,\underline{v} + \underline{u})'] \\ &= E[(\underline{N}\,\underline{v} + \underline{u})\,(\underline{v}'\,\underline{N}' + \underline{u}')] \\ &= E[\underline{N}\,\underline{v}\,\underline{v}'\,\underline{N}' + \underline{u}\,\underline{v}'\,\underline{N}' + \underline{N}\,\underline{v}\,\underline{u}' + \underline{u}\,\underline{u}'] \\ &= \underline{N}\,E(\underline{v}\,\underline{v}')\underline{N}' + E(\underline{u}\,\underline{v}')\underline{N}' + \underline{N}\,E(\underline{v}\,\underline{u}') + E(\underline{u}\,\underline{u}') \\ &= \underline{N}\,\underline{\Omega}\,\underline{N}' + \underline{\Sigma} = \underline{\Theta} \end{aligned}$$

6.3. The OLS Estimator

As described in the previous section, when the transition probabilities are assumed to be variable over time, varying with exogenous explanatory variables, the model can be written in the form of equations (6.3) and (6.4)

$$\underline{n}^* = \underline{N}\ \underline{p} + \underline{u}$$
$$\underline{p} = \underline{X}\ \underline{\delta} + \underline{v}$$

or, in reduced form, as equations (6.5) and (6.6)

$$\underline{n}^* = \underline{N}\ \underline{X}\ \underline{\delta} + \underline{w}$$
$$\underline{w} = \underline{N}\ \underline{v} + \underline{u}$$

where

$$E(\underline{u}) = \underline{0} \qquad E(\underline{v}) = \underline{0} \qquad E(\underline{u}\ \underline{v}) = \underline{0}$$
$$E(\underline{u}\ \underline{u}') = \underline{\Sigma} \qquad E(\underline{v}\ \underline{v}') = \underline{\Omega} \qquad E(\underline{w}) = \underline{0}$$
$$E(\underline{w}\ \underline{w}') = \underline{N}\ \underline{\Omega}\ \underline{N}' + \underline{\Sigma} = \underline{\Theta}$$

The unrestricted OLS estimation procedure can be applied to equations (6.5) in order to get the first estimates for the parameters $\underline{\delta}$. Thus, the problem is to find the estimate $\hat{\underline{\delta}}^s$ that minimises the sum of the squared residuals

$$\underline{w}'\underline{w} = (\underline{n}^* - \underline{N}\ \underline{X}\ \hat{\underline{\delta}})'\ (\underline{n}^* - \underline{N}\ \underline{X}\ \hat{\underline{\delta}})$$

The right-hand side of the equation can be written as

$$\begin{aligned}(\underline{n}^* - \underline{N}\ \underline{X}\ \hat{\underline{\delta}})'\ (\underline{n}^* - \underline{N}\ \underline{X}\ \hat{\underline{\delta}}) &= (\underline{n}^{*\prime} - \hat{\underline{\delta}}'\underline{X}'\underline{N}')(\underline{n}^* - \underline{N}\ \underline{X}\ \hat{\underline{\delta}}) \\ &= \underline{n}^{*\prime}\underline{n}^* - \underline{n}^{*\prime}\underline{N}\ \underline{X}\ \hat{\underline{\delta}} - \hat{\underline{\delta}}'\underline{X}'\underline{N}'\underline{n}^* \\ &\quad + \hat{\underline{\delta}}'\underline{X}'\underline{N}'\underline{N}\ \underline{X}\ \hat{\underline{\delta}} \\ &= \underline{n}^{*\prime}\underline{n}^* - 2\ \hat{\underline{\delta}}'\underline{X}'\underline{N}'\underline{n}^* + \hat{\underline{\delta}}'\underline{X}'\underline{N}'\underline{N}\ \underline{X}\ \hat{\underline{\delta}}\end{aligned}$$

The minimisation of the error sum of squares is given by

$$\frac{\partial(\underline{w}'\underline{w})}{\partial\ \hat{\underline{\delta}}} = -2\ \underline{X}'\underline{N}'\underline{n}^* + 2\ \underline{X}'\underline{N}'\underline{N}\ \underline{X}\ \hat{\underline{\delta}} = \underline{0}$$

that is,

$$\underline{X}'\underline{N}'\underline{n}^* = \underline{X}'\underline{N}'\underline{N}\ \underline{X}\ \hat{\underline{\delta}}$$

$$\hat{\underline{\delta}} = (\underline{X}'\underline{N}'\underline{N}\ \underline{X})^{-1}\underline{X}'\underline{N}'\underline{n}^*$$

or

$$\hat{\underline{\delta}} = [(\underline{N}\ \underline{X})'(\underline{N}\ \underline{X})]^{-1}\ (\underline{N}\ \underline{X})'\ \underline{n}^* \qquad (6.7)$$

where $(\underline{N}\ \underline{X})'(\underline{N}\ \underline{X})$ is a non-singular symmetric block diagonal matrix, providing that $k > (s+1)(m+1)$. The estimated vector for the transition probabilities is then determined by

$$\hat{\underline{p}} = \underline{X}\ \hat{\underline{\delta}} \qquad (6.8)$$

which means that each estimate of a transition probability is a linear combination of the estimates of the parameters $\underline{\delta}$. The expression for each estimate is then

$$\hat{p}_{ij}(t) = \hat{a}_{ij} + \sum_{h=1}^{m} x_h(t)\ \hat{\delta}_{ijh} \tag{6.9}$$

Having obtained the elementary statistics (mean value, variance and covariance) for the parameters $\hat{\delta}$, the corresponding mean value and variance for each transition probability's estimate can be calculated, using equation (6.9), as follows

$$\text{(i)}\quad E[\hat{p}_{ij}(t)] = \sum_{h=1}^{m} x_h(t)\ E(\hat{\delta}_{ijh})$$

$$\begin{aligned}\text{(ii)}\quad \mathrm{var}[\hat{p}_{ij}(t)] &= \mathrm{var}\Big[\sum_{h=1}^{m} x_h(t)\ \hat{\delta}_{ijh}\Big] \\ &= \sum_{h=1}^{m} x_h^2(t)\ \mathrm{var}(\hat{\delta}_{ijh}) \\ &\quad + 2\sum_{h=1}^{m-1}\ \sum_{g=h+1}^{m} x_h(t)\ x_g(t)\ \mathrm{cov}(\hat{\delta}_{ijh}, \hat{\delta}_{ijg})\end{aligned} \tag{6.10}$$

The unrestricted OLS estimator does not necessarily satisfy the non-negativity and row-sum conditions. If a restricted OLS estimation procedure is performed, a quadratic programming problem arises. The problem becomes as follows

- find the estimate $\hat{\underline{\delta}}^{s}$ which minimises the positive quadratic form

$$\Phi = \underline{w}'\underline{w} = (\underline{n}^* - \underline{N}\ \underline{X}\ \hat{\underline{\delta}})'\ (\underline{n}^* - \underline{N}\ \underline{X}\ \hat{\underline{\delta}})$$

subject to

$$\underline{G}\ \underline{X}\ \underline{\hat{\delta}} = \underline{\eta}_{(s+1)k}$$

$$\underline{X}\ \underline{\hat{\delta}} \geq \underline{0}$$

where $\underline{G}$ is a $[(s+1)k \times (s+1)^2k]$ matrix of type $[A_1, \ldots, A_{s+1}]$ and each $\underline{A_i}$ is a $[(s+1)k \times (s+1)k]$ diagonal matrix of entries zero or unity in the main diagonal. The vector $\underline{\eta}_{(s+1),k}$ is a unit vector of size $[(s+1)k \times 1]$. The SYMQUAD algorithm can be employed to compute the optimum value of $\underline{\hat{\delta}}$ (see D.1, Appendix D). The restricted OLS estimates of the transition probabilities can then be obtained using equation (6.8)

$$\underline{\hat{p}} = \underline{X}\ \underline{\hat{\delta}}^{s}$$

and the statistics can be obtained following the same process described for the unrestricted OLS, that is, using equations (6.10).

6.4 The GLS Estimator

The estimates of the transition probabilities derived by solving the OLS estimation procedure, either unrestricted or restricted, assume that the covariance matrices $\underline{\Sigma}$, $\underline{\Omega}$ and $\underline{\Theta}$ are of the form of a positive scalar times identity matrices. The variances of the disturbance terms may not be constant and the OLS estimates obtained are not efficient and heteroscedasticity arises. The problem of heteroscedasticity can be overcome by applying the GLS estimation procedure. Also, if s of the equations of the model are known, the remaining equation that gives the number of students who leave the school system, through the drop-out probabilities, is therefore determined by the row sum condition. Being more interested in estimating the repetition and promotion probabilities, the last equation of the model may be deleted. The reduced form of the model can then be written as

$$\underline{n} = \underline{N}\ \underline{X}\ \underline{\delta} + \underline{w} \qquad (6.11)$$

$$\underline{w} = \underline{N}\ \underline{v} + \underline{u}$$

ith

$$E\ (\underline{w}) = \underline{0}$$

$$E\ (\underline{w}\ \underline{w}') = \underline{\Theta}$$

where $\underline{\Theta}$ is a (sk x sk) non-singular matrix. In a more detailed form, equation (6.11) can be described as

$$\begin{bmatrix} \underline{n}_1^* \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \underline{n}_s^* \end{bmatrix} = \begin{bmatrix} \underline{N}_1 & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{N}_2 & & \underline{0} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \underline{0} & \underline{0} & & \underline{N}_s \end{bmatrix} \begin{bmatrix} \underline{X}_1 & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \underline{X}_2 & & \underline{0} \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot \\ \underline{0} & \underline{0} & & \underline{X}_s{}^2 \end{bmatrix} \begin{bmatrix} \underline{\delta}_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \underline{\delta}_s \end{bmatrix} + \begin{bmatrix} \underline{w}_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \underline{w}_s \end{bmatrix}$$

Proceeding in the same way as in the case of the basic stationary model, the unrestricted GLS estimator for $\underline{\delta}$ is

$$\hat{\underline{\delta}} = [(\underline{N}\ \underline{X})'\ \underline{\Theta}^{-1}(\underline{N}\ \underline{X})]^{-1}(\underline{N}\ \underline{X})'\underline{\Theta}^{-1}\ \underline{n}^*$$

and equation (6.9) applied to the new estimates $\hat{\underline{\delta}}$ gives the values of the new estimates $\hat{\underline{p}}$ for the transition probabilities. The deleted parameters $\underline{p}_{s+1}$ may be estimated using the expression

$$\hat{\underline{p}}_{s+1} = \underline{n}_{(s+1)k} - \underline{R}\,\hat{\underline{p}}$$

where $\underline{R}$ is a submatrix of $\underline{G}$, with the form $[\underline{A}_1, \ldots , \underline{A}_s]$ with each $\underline{A}_i$ a matrix of size $[(s+1)k \times (s+1)k]$, and $\underline{n}_{(s+1)k}$ the $[(s+1)k \times 1]$ unit vector.

The restricted GLS estimator is obtained after including the non-negativity and row-sum conditions in the model. This leads again to a quadratic programming problem that now has the following form

- find $\hat{\underline{\delta}}^s$ which minimises the positive quadratic form

$$\underline{\Phi} = (\underline{n}^* - \underline{N}\ \underline{X}\ \hat{\underline{\delta}})'\ \underline{\Theta}^{-1}(\underline{n}^* - \underline{N}\ \underline{X}\ \hat{\underline{\delta}})$$

subject to

$$\underline{R}\ \underline{X}\ \hat{\underline{\delta}} = \underline{n}_{(s+1)k}$$
$$\underline{X}\ \hat{\underline{\delta}} \geq \underline{0}$$

The algorithm described in D.2, Appendix D, gives the estimates of $\hat{\underline{\delta}}$. Using equations (6.9), the restricted GLS estimates $\hat{\underline{p}}$ for the transition probabilities are then obtained.

The basic operational difficulty when applying the GLS estimator, either unrestricted or restricted, is that the matrices $\underline{\Sigma}$ and $\underline{\Omega}$ are unknown and so the matrix $\underline{\Theta}$ is also unknown. However, the estimates of the elements of this last matrix can be obtained following the process proposed by ZELLNER [1962] and already referred to in Chapter 4 for the basic model. Matrix $\underline{\Theta}$ can be written as

$$\underline{\Theta} = \underline{N}\,\underline{\Omega}\,\underline{N}' + \underline{\Sigma} = E(\underline{w}\,\underline{w}') = \begin{bmatrix} E(\underline{w}_1\,\underline{w}_1') & \cdots & E(\underline{w}_1\,\underline{w}_s') \\ \cdots & \cdots & \cdots \\ E(\underline{w}_s\,\underline{w}_1') & \cdots & E(\underline{w}_s\,\underline{w}_s') \end{bmatrix}$$

where each $E(\underline{w}_i\,\underline{w}_i')$ is a (k x k) variance - covariance matrix of the disturbances in the ith equation of (6.5). By assumption, $E(\underline{w}_i\,\underline{w}_j') = \sigma_{ij}\,\underline{I}$, with i,j = 1, ... , s, which means that the disturbance in any of the s equations of (6.11) is homoscedastic and non-autocorrelated. Matrix $\underline{\Theta}$ can then be written as

$$\underline{\Theta} = \underline{N}\,\underline{\Omega}\,\underline{N}' + \underline{\Sigma} = \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1s} \\ \sigma_{21} & \sigma_{22} & \cdots & \sigma_{2s} \\ \cdots & \cdots & \cdots & \cdots \\ \sigma_{s1} & \sigma_{s2} & \cdots & \sigma_{ss} \end{bmatrix} \otimes \underline{I}$$

where I is a unit matrix of order (k x k) and $\otimes$ denotes the Kronecker multiplication of matrices. Thus σ_{ij} are estimated by

$$s_{ij} = (\underline{n}_i^* - \underline{N}_i\,\underline{X}_i\,\hat{\underline{\delta}}_i)'\,(\underline{n}_j^* - \underline{N}_j\,\underline{X}_j\,\hat{\underline{\delta}}_j)/(k-m)$$

$$i,j = 1, \ldots, s$$

Chapter 6

Footnotes

1. The unavailability for most of the situations of micro data (that is students in the above study), describing how individual units have behaved through time, leads to the use of aggregate time-series of the number of the units in each state at time t. If micro data were available, Maximum Likelihood techniques would be more appropriate [see LEE, JUDGE and ZELLNER, 1970].

Chapter 7

APPLICATION OF THE TIME VARYING MARKOV MODEL TO THE PORTUGUESE EDUCATIONAL SYSTEM

7.1. Introduction

This chapter will concentrate on the application of the extended Markov model developed in the previous chapter to the same sectors of the Portuguese educational system already studied in Chapters 4 and 5, using the country as a whole.

Seventeen explanatory variables, broadly divided into supply factors and demand factors, are included in the extended model and the unrestricted and restricted OLS estimation procedure will be performed in order to obtain time varying estimates for the transition probabilities (repetitions and promotions).

The existence of multicollinearity between the explanatory variables suggests that a principal component analysis should be attempted. Regression on the principal components of the explanatory variables is therefore performed and new estimates for the transition probabilities are obtained.

A summary description of the programs used in this chapter and displayed in Appendix E is presented in Appendix G.

7.2. Selection of the Explanatory Variables

Different sources were thoroughly searched in order to gather the data for the explanatory variables. The impossibility of using cross-section data describing students' previous school performance, their attitudes and aspirations, and also the social, educational and professional status of the parents, has confined this study to the use of aggregate data. Furthermore, even within the aggregate data, the type of information needed for the work was not always available. Specific variables for which data were available, and which appear to be relevant to the present study, have therefore been selected.. The choice of these variables follows the recomendations published and take into account the results of the previous studies referred in the literature review.

The list of explanatory variables selected, and whose impact on the transition probabilities this study attempts to test, totals seventeen. For an easier understanding of causality, these determinants were broadly divided into those belonging to the supply side and those belonging to the demand side of the education system. For the supply side, the selected variables have the aim of specifying changes in the quality of school places, school characteristics affecting the students' behaviour (e.g. number and qualification of teachers), social facilities offered to the students, etc. For the demand side,[1] the selected variables may influence the students' demand for education, such as future earnings prospects and job opportunities.

The list of the explanatory variables used in this study is shown in Table 7.1. Details of the observed values of these

Table 7.1 : Original Explanatory Variables

Variable	Code name	Description
Supply Factors		
X1	EDUC*	percentage increase in education expenditures.
X2	PEDUC	percentage of education expenditures in GDP.
X3	PCAP	percentage of capital expenditures in education expenditures
X4	COST*	private expenditure per student.
X5	TEARN*	percentage increase in teachers' earnings
X6	PUTEA	pupil-teacher ratio.
X7	PUCLASS	pupil-classroom ratio.
X8	BUS	percentage of students who use school bussing
X9	HELP	percentage of students who get scholarship from the social services
X10	UNQUAL	percentage of teachers without qualifications.
Demand Factors		
X11	GDP*	gross domestic product.
X12	LIFE	life expectation at birth.
X13	ILLIT	illiteracy rate.
X14	UNEMP	number of unemployed workers.
X15	EARN*	monthly worker salaries (mean values)
X16	LFLEV	percentage of labour force with a preparatory or secondary level of education.
X17	POPLEV	percentage of population with a preparatory or secondary level of education.

* values referred to 1970 prices

TABLE 7.2: Original Explanatory Variables for the Supply-Side - Whole Country

YEAR	EDUC	PEDUC	PCAP	COST[a]	TEARN[a]	PUTEA[a]	PUCLASS[a]	BUS[a]	HELP[a]	UNQUAL[a]
1971	4.6	2.1	19.1	4.8	1.5	18.8	38.0	0.0	0.0	51.9
1972	18.0	2.3	16.1	5.1	- 0.6	17.2	51.0	0.0	0.0	42.6
1973	22.7	2.7	20.2	2.9	- 3.4	16.8	47.8	3.5	12.8	35.5
1974	1.5	2.8	13.9	5.1	-10.7	16.5	52.6	5.6	13.8	30.4
1975	45.8	4.3	11.7	7.0	34.6	14.1	51.0	9.5	16.2	25.8
1976	3.3	4.3	13.8	6.8	8.7	14.5	48.7	16.3	21.7	22.3
1977	14.4	4.6	11.0	6.7	-12.1	13.7	47.9	14.8	15.9	19.5
1978	1.3	4.1	13.9	6.5	4.8	14.3	47.7	17.8	15.9	17.4
1979	7.6	3.8	11.9	6.8	6.8	13.2	47.3	19.7	16.1	15.9
1980	22.7	4.3	14.6	8.2	5.4	12.4	46.8	25.5	20.0	13.5
1981	2.9	4.4	13.6	8.4	3.1	12.2	45.3	28.0	24.4	16.4
$\bar{x}$	13.16	3.61	14.53	6.21	3.46	14.88	47.65	12.79	14.25	26.47
sd	13.620	0.938	2.917	1.607	12.345	2.140	3.840	9.788	7.830	12.380

a) Weighted mean of the observed values for the preparatory plus secondary levels.

TABLE 7.3: Original Explanatory Variables for the Demand-Side - Whole Country

YEAR	GDP (10^9)	LIFE	ILLIT	UNEMP	EARN	LFLEV	POPLEV
1971	177.5	67.16	24.9	52955	1947	10.0	13.2
1972	186.7	67.56	24.2	69244	1947	11.5	13.5
1973	201.3	68.02	23.5	70433	1947	12.3	13.9
1974	193.4	68.55	22.8	112682	3071	13.3	13.4
1975	186.5	69.14	22.2	140922	3156	14.7	14.4
1976	191.4	69.20	21.5	210053	2910	15.4	15.4
1977	205.9	70.52	20.8	257871	2698	17.0	16.4
1978	225.6	71.30	20.1	304388	2650	18.5	17.7
1979	253.1	72.15	19.6	308968	2524	19.6	19.4
1980	275.5	73.06	18.9	269290	2615	23.4	20.3
1981	275.2	74.04	18.3	244470	2677	25.4	22.0
$\bar{x}$	215.46	70.06	21.53	185573	2558	16.49	16.33
sd	35.950	2.316	2.193	98652	427	4.957	3.100

variables are given in Table 7.2 and Table 7.3 and the corresponding standardised values are presented in Table 7.4 and Table 7.5.

As the precision of estimating falls if there is multicollinearity between the 'independent' variables of the model, a better understanding of the interrelationships between the selected explanatory variables to be used in the extended Markov model, was sought. The correlation matrix was computed, and is presented in Table 7.6. It shows that for most of the cases, high values for the correlation coefficients between the variables have been found. This phenomenon not only occurs between two supply factors or between two demand factors, but also between supply-side and demand-side side variables.

Using the F-statistic with n = 11 and k = 2, the values for the correlation coefficient R under which it is acceptable to assume no dependence between the variables, is 0.7245 at the 1% significance level and 0.5961 at the 5% significance level. The great number of values over the maximum acceptable correlation coefficient, even at the 1% significance level, and indicated with an * in Table 7.6, confirm the existence of multicollinearity in the set of explanatory variables selected.

Instead of, at this stage, dropping variables that may have some appreciable effect in the change of the transition probabilities, it seems worthwhile to use all the seventeen variables and to perform stepwise regression, or regression on the principal components of these explanatory variables.

TABLE 7.4: Standardised Values for the Supply Side Explanatory Variables - Whole Country

YEAR	EDUC	PEDUC	PCAP	COST	TEARN	PUTEA	PUCLASS	BUS	HELP	UNQUAL
1971	-2.09	-5.34	5.21	-2.92	-0.53	6.07	-8.31	-4.34	-6.04	6.80
1972	1.16	-4.63	1.79	-2.29	-0.96	3.55	2.90	-4.34	-6.04	4.35
1973	2.32	-3.20	6.43	-6.83	-1.86	2.98	0.13	-3.15	-0.63	2.42
1974	-2.85	-2.85	-0.73	-2.29	-3.81	2.53	4.28	-2.42	-0.20	1.06
1975	7.96	2.45	-3.21	1.63	8.36	-1.19	2.89	-1.13	0.83	-0.17
1976	-2.38	2.45	-0.83	1.23	1.39	-0.60	0.90	1.19	3.15	-1.16
1977	0.30	3.52	-3.98	1.03	-4.18	-1.82	0.23	0.70	0.70	-1.86
1978	-2.88	1.67	-0.73	0.60	0.33	-0.90	0.03	1.69	0.70	-2.45
1979	-1.35	0.67	-2.98	1.23	0.86	-2.62	-0.30	2.35	0.76	-2.82
1980	2.32	2.45	0.07	4.10	0.50	-3.85	-0.73	4.31	2.45	-3.48
1981	-2.49	2.79	-1.02	4.51	-0.10	-4.15	-2.02	5.14	4.32	-2.69

TABLE 7.5: Standardised Values for the Demand Side Explanatory Variables - Whole Country

YEAR	GDP	LIFE	ILLIT	UNEMP	EARN	LFLEV	POPLEV
1971	-3.52	-4.14	5.55	-4.45	-4.64	-4.34	-3.35
1972	-2.65	-3.58	4.05	-3.92	-4.64	-3.35	-3.02
1973	-1.29	-2.92	2.98	-3.88	-4.64	-2.82	-2.59
1974	-2.02	-2.16	1.92	-2.45	3.88	-2.12	-3.15
1975	-2.69	-1.33	1.03	-1.49	4.54	-1.19	-2.06
1976	-2.22	-1.23	-0.03	0.83	2.68	-0.73	-0.99
1977	-0.90	0.66	-1.09	2.42	1.06	0.32	0.07
1978	0.93	1.73	-2.19	3.98	0.70	1.36	1.46
1979	3.48	2.99	-2.92	4.15	-0.27	2.09	3.28
1980	5.37	4.31	-3.98	2.82	0.43	4.61	4.25
1981	5.51	5.67	-4.88	1.99	0.90	6.17	6.10

TABLE 7.6: Correlation Matrix for the Explanatory Variables

	EDUC	PEDUC	PCAP	COST	TEARN	PUTEA	PUCLASS	BUS	HELP	UNQUAL
EDUC	1.00000	.13174	- .07733	- .00093	.62714	- .09115	.30713	- .17045	- .03374	.03355
PEDUC		1.00000	- .77041*	.81610*	.32295	- .91792*	.19103	.83890*	.86095*	- .90033*
PCAP			1.00000	.73315*	-.25303	.72673*	- .44178	- .56837	- .55872	.72478
COST				1.00000	.35539	- .86388*	.00549	.86178*	.66351	- .76034*
TEARN					1.00000	- .26530	.08506	.13766	.19145	- .14816
PUTEA						1.00000	- .19728	.94255*	.86131*	.95581*
PUCLASS							1.00000	- .03086	.24152	- .29437
BUS								1.00000	.85319*	- .91408*
HELP									1.00000	- .89035*
UNQUAL										1.00000
GDP										
LIFE										
ILLIT										
UNEMP										
EARN										
LFLEV										
POPLEV										

* - Values over the maximum acceptable correlation coefficient at 1% significance level

TABLE 7.6: Correlation Matrix for the Explanatory Variables (Continued)

	GDP	LIFE	ILLIT	UNEMP	EARN	LFLEV	POPLEV
EDUC	- 26182	- 21670	.19232	- .15254	.07443	- .24204	- .24497
PEDUC	.58787	.77682*	- .85975*	.87207*	.66876	.76641*	.71903
PCAP	-.27251	-.52758	.61970	- .69387	- .81164*	- .51801	- .41006
COST	.58946	.77140*	- .79262*	.79524*	.57136	.75048*	.74567*
TEARN	-.11551	.02100	- .04542	- .00340	.36752	.04025	- .00615
PUTEA	-.68745	-.85449*	.90795*	- .89047*	- .57324	- .83141*	- .80271*
PUCLASS	-.09137	-.03329	- .06881	.03929	.50066	- .00659	- .14358
BUS	.79020*	.91275*	- .93895*	.89543*	.46130	.89132*	.89307*
HELP	.59185	.73496*	- .81776*	.72831*	.68294	.74007*	.67366
UNQUAL	-.69824	-.84538*	.92130*	- .91648*	- .63097	- .82361*	- .77201
GDP	1.00000	.94830*	- .89452	.72884*	.13277	.94275*	.96427*
LIFE		1.00000	- .98254*	.86086*	.33399	.99231*	.98491*
ILLIT			1.00000	- .91535*	- .47427	- .97403*	- .94941*
UNEMP				1.00000	.45416	.81721*	.82111*
EARN					1.00000	.37613	.22526
LFLEV						1.00000	.97883*
POPLEV							1.00000

7.3. The Extended Model Applied to the Educational System

The application of a Markov model to the educational system allows some simplifications with reference to the transition probability matrix $\underline{P}^t_{a,t-1}$. The movements of the students within the educational system are always from one state (in this study, grade) of the system to another state. Furthermore, a student cannot skip two states at a time. Thus, the transition probability matrix is an upper triangular matrix, with all the elements p_{ij} of the matrix, $i > j$, being equal to zero. Also, some elements of the upper triangle of the transition probability matrix are equal to zero. In order to increase the number of the degrees of freedom and reduce the size of the matrices involved, all of those zero elements can be eliminated from the equations. Therefore the equations of the model, for each grade and for each academic year, can be written as follows:[2]

$$
\begin{aligned}
n_1^t &= n_1^{t-1}\; p_{11}^t + u_1^t \\
n_i^t &= n_{i-1}^{t-1}\; p_{i-1,i}^t + n_i^{t-1}\; p_{ii}^t + u_i^t \qquad\qquad i=2, \ldots, s \\
n_{s+1}^t &= \sum_{i=1}^{s} n_i^{t-1}\; p_{i,s+1}^t + u_{s+1}^t \qquad\qquad t=1, \ldots, k
\end{aligned}
\tag{7.1}
$$

For simplicity it has been assumed that the equations describing the relationship between the explanatory variables and the transition probabilities have a constant term associated with each probability, the explanatory variables affecting all the repetition probabilities, all the promotion probabilities and all the drop-out probabilities in the same way, respectively.

Furthermore, the parameters $\underline{\delta}$ were classified into three groups $\underline{\delta}_r$, $\underline{\delta}_p$ and $\underline{\delta}_d$, each being associated with the repetition probability, the promotion probability and the drop-out probability, respectively. The equations that describe these relationships are

$$
\begin{aligned}
p_{ii}^{t} &= a_{ii} + \underline{x}^{t'}\, \underline{\delta}_r && i=1, \ldots ,s \\
p_{i-1,i}^{t} &= a_{i-1,i} + \underline{x}^{t'}\, \underline{\delta}_p && i=2, \ldots ,s \qquad (7.2) \\
p_{i,s+1}^{t} &= a_{i,s+1} + \underline{x}^{t'}\, \underline{\delta}_d && i=1, \ldots ,s \\
& && t=1, \ldots ,k
\end{aligned}
$$

Thus, using equations (7.1) and equations (7.2), the extended model can be written as:

$$n_1^t = a_{11}\, n_1^{t-1} + n_1^{t-1}\, \underline{x}^{t'}\, \underline{\delta}_r + u_1^t$$

$$n_i^t = (a_{i-1,i} + \underline{x}^{t'}\, \underline{\delta}_p)\, n_{i-1}^{t-1} + (a_{ii} + \underline{x}^{t'}\, \underline{\delta}_r)\, n_i^{t-1} + u_i^t \qquad i=2, \ldots ,s$$

$$n_{s+1}^t = \sum_{i=1}^{s} a_{i,s+1}\, n_i^{t-1} + \sum_{i=1}^{s} n_i^{t-1}\, \underline{x}^{t'}\, \underline{\delta}_d + u_{s+1}^t$$

That is,

$$n_1^t = a_{11}\, n_1^{t-1} + n_1^{t-1}\, \underline{x}^{t'}\, \underline{\delta}_r + u_1^t$$

$$n_i^t = a_{i-1,i}\, n_{i-1}^{t-1} + a_{ii}\, n_i^{t-1} + (n_i^{t-1}\, \underline{x}^{t'},\ n_{i-1}^{t-1}\, \underline{x}^{t'}) \begin{bmatrix} \underline{\delta}_r \\ \underline{\delta}_p \end{bmatrix} + u_i^t$$

$$i=2, \ldots, s$$

$$n_{s+1}^t = \sum_{i=1}^{s} a_{i,s+1}\, n_i^{t-1} + \sum_{i=1}^{s} n_i^{t-1}\, \underline{x}^{t'}\, \underline{\delta}_d + u_{s+1}^t \qquad (7.3)$$

$$t=1, \ldots, k$$

where $\underline{x}^{t'}$ is the (1 x m) row vector of the explanatory variables and $\underline{\delta}_r$, $\underline{\delta}_p$, and $\underline{\delta}_d$ are the (m x 1) column vectors of the corresponding coefficients. The number of equations in (7.3) is (s+1)k. Stacking over time and over variables, these equations can be set out compactly in matrix notation as

$$\underline{n}^* = \underline{N}^x\, \underline{\delta} + \underline{u} \qquad (7.4)$$

where $\underline{n}^*$ is the [(s+1)k x 1] vector of the number of students by grade, $\underline{u}$ is the [(s+1)k x 1] vector of disturbance terms, $\underline{\delta}$ is the [(3s-1+3m) x 1] vector of parameters of the form $\underline{\delta}' = (\underline{a}', \underline{\delta}_r', \underline{\delta}_p', \underline{\delta}_d')$, $\underline{a}$ being a vector of size [(3s-1) x 1] and $\underline{\delta}_r$, $\underline{\delta}_p$, $\underline{\delta}_d$ being vectors of size (m x 1). These vectors have the following forms:

$$\underline{n}^* = \begin{bmatrix} \underline{n}^{*t}_1 \\ \underline{n}^{*t}_2 \\ \cdot \\ \cdot \\ \cdot \\ \underline{n}^{*t}_{s+1} \end{bmatrix} \qquad \underline{u} = \begin{bmatrix} \underline{u}_1^t \\ \underline{u}_2^t \\ \cdot \\ \cdot \\ \cdot \\ \underline{u}_{s+1}^t \end{bmatrix}$$

$$\underline{a} = \begin{bmatrix} a_{11} \\ a_{12} \\ a_{22} \\ a_{23} \\ \cdot \\ \cdot \\ \cdot \\ a_{ss} \\ a_{1,s+1} \\ a_{2,s+1} \\ \cdot \\ \cdot \\ \cdot \\ a_{s,s+1} \end{bmatrix} \qquad \underline{\delta}_r = \begin{bmatrix} \delta_{r1} \\ \delta_{r2} \\ \cdot \\ \cdot \\ \cdot \\ \delta_{rm} \end{bmatrix} \qquad \underline{\delta}_p = \begin{bmatrix} \delta_{p1} \\ \delta_{p2} \\ \cdot \\ \cdot \\ \cdot \\ \delta_{pm} \end{bmatrix} \qquad \underline{\delta}_d = \begin{bmatrix} \delta_{d1} \\ \delta_{d2} \\ \cdot \\ \cdot \\ \cdot \\ \delta_{dm} \end{bmatrix}$$

Matrix $\underline{N}^x$ is the $[(s+1)k \times (3s-1+3m)]$ block matrix incorporating the number of students by grade in the previous year and the exogenous explanatory variables. The matrix $\underline{N}^x$ can be expressed by two block matrices as follows

$$\underline{N}^x = [\ \underline{N}^o \ \vdots \ \underline{n}^{t-1} \otimes \underline{x}^t\ ]$$

where the left-hand side matrix has the form

$$\underline{N}^o = \begin{bmatrix}
\underline{n}_1^{t-1} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \cdots & \underline{0} & \underline{0} & \underline{0} \cdots & \underline{0} \\
\underline{0} & \underline{n}_1^{t-1} & \underline{n}_2^{t-1} & \underline{0} & \underline{0} & & \cdot & \cdot & \cdot & \cdot \\
\cdot & \underline{0} & \underline{0} & \underline{n}_2^{t-1} & \underline{n}_3^{t-1} & & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \underline{0} & \underline{0} & & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & & \underline{0} & \underline{0} & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & & \underline{n}_{s-1}^{t-1} & \underline{n}_s^{t-1} & \underline{0} & \underline{0} \\
\underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \cdots & \underline{0} & \underline{0} & \underline{n}_1^{t-1} \cdots & \underline{n}_s^{t-1}
\end{bmatrix}$$

and the right-hand side matrix is

$$\underline{n}^{t-1} \otimes \underline{x}^{t} = \begin{bmatrix} \underline{n}_1^{t-1} \otimes \underline{x}^t & \underline{0} & \underline{0} \\ \underline{n}_2^{t-1} \otimes \underline{x}^t & \underline{n}_1^{t-1} \otimes \underline{x}^t & \underline{0} \\ \underline{n}_3^{t-1} \otimes \underline{x}^t & \underline{n}_2^{t-1} \otimes \underline{x}^t & \underline{0} \\ \vdots & \vdots & \vdots \\ \underline{n}_s^{t-1} \otimes \underline{x}^t & \underline{n}_{s-1}^{t-1} \otimes \underline{x}^t & \underline{0} \\ \underline{0} & \underline{0} & \sum_{i=1}^{s} \underline{n}_i^{t-1} \otimes \underline{x}^t \end{bmatrix}$$

where $\otimes$ represents the Kronecher product and $\underline{x}^t$ is the (k x m) matrix $\underline{x}^t = [\underline{x}_1^t \; \ldots \; \underline{x}_m^t]$ of explanatory variables.

When applying this model to the case study of the Portuguese educational system, the dimensions of the variables are k=11, s=7, m=17. Therefore, vectors $\underline{n}^*$ and $\underline{u}$ have size (88 x 1), matrix $\underline{N}^x$ has size (88 x 71) and vector $\underline{\delta}$ has size (71 x 1).

7.4. The OLS Estimator

7.4.1. The Unrestricted OLS Estimator

As shown in the previous section, the extended Markov model applied to the educational system can be described, in compact form, by equation (7.4)

$$\underline{n}^* = \underline{N}^x \, \underline{\delta} + \underline{u}$$

The unrestricted OLS estimator is then given by the following expression

$$\underline{\hat{\delta}} = (\underline{N}^{x'}\underline{N}^{x})^{-1}\ \underline{N}^{x'}\ \underline{n}^{*} \qquad (7.5)$$

It has been assumed that the explanatory variables effect only the changes in the repetition and promotion probabilities, the drop-out probabilities being estimated using the expression

$$\hat{p}^{t}_{i,s+1} = 1 - \hat{p}^{t}_{ii} - \hat{p}^{t}_{i,i+1}$$

Thus the last set of equations was removed from the model, thereby increasing the number of degrees of freedom. Matrix $\underline{N}^{x}$ has then size (77 x 47), in which $\underline{N}^{o}$ has size (77 x 13) and $\underline{n}^{t-1} \otimes \underline{x}^{t}$ has size (77 x 34); $\underline{\delta}$ is a vector of size (47 x 1) and $\underline{n}^{*}$ is a vector of size (77 x 1). The Fortran program PROGRM(VAR) and the job program JOBB (see Appendix E) have been used to generate the data matrix $[\underline{n}^{*} \vdots \underline{N}^{x}]$ of size (77 x 48), the data file for the regression problem. The subprogram REGRESSION included in the SPSS version 7 programs available at the University of Manchester Regional Computer Centre was then applied to this data file (Method I) in order to get the coefficients $\underline{\hat{a}}$, $\underline{\hat{\delta}}_{r}$ and $\underline{\hat{\delta}}_{p}$ for this model. However, due to the existence of multicollinearity between the explanatory variables used in this study, the stepwise method has been selected. In this case, the stepwise estimators are preferred to the OLS estimators because they have smaller mean square errors [see WALLACE, 1964]. With this method, the variables corresponding to the $\underline{a}$ coefficients have been forced into the equation, the stepwise method being applied to the remaining variables. Any variable with a F-ratio less than a predetermined minimum value (.005) was considered eligible for removal.

Program REG(EXTCON) presented in Appendix E was used and the results are shown in Table 7.7 and Table 7.8. For an easier reading of these tables, coefficients $\underline{a}$ are assigned as a_{55} to $a_{11,11}$, where each coefficient is the constant term associated with each type of transition probability p_{55} to $p_{11,11}$(see equations 7.2). In Table 7.8 the listed variables are followed by a number, one or two. The coefficients associated with variables <name>1 refer to the elements of the $\underline{\delta}_r$ vector and the coefficients associated with variables <name>2 refer to the elements of the $\underline{\delta}_p$ vector. This means that variables PEDUC1 to POPLEV1 are the variables selected to determine the changes in the repetition probabilities, and variables PCAP1 to POPLEV2 are the variables which determine the changes in the promotion probabilities. Sixteen variables have been removed from the equation, eight of them corresponding to the repetition probabilities and the other eight corresponding to the promotion probabilities. Within these removed variables, five (EDUC, COST, UNQUAL, ILLIT, LFLEV) have been removed from both types of equation. From Table 7.6, however, it can be seen that although variable EDUC is not correlated with any other variable, the same is not true for the other four variables. Thus, it seems reasonable to think that EDUC is the only variable to have no effect when the unrestricted OLS regression method is performed. The remaining variables are highly correlated with variables that stay in the equation; some of the relationships are described, by the following expressions

$$\text{COST} = 0.00079 - \underset{(5.15)}{0.86700}\ \text{PUTEA} \qquad R = .86388$$

$$\text{UNQUAL} = -0.00087 + \underset{(9.75)}{0.96029}\ \text{PUTEA} \qquad R = .95581$$

$$\text{LFLEV} = 0.00000 + \underset{(14.34)}{0.97792}\ \text{POPLEV} \qquad R = .97883$$

Table 7.7 : Estimate of the $\underline{a}$ Coefficients for the Unrestricted OLS for the Whole Country (Method I)

Coefficient	Value	St.error	t-value	95% Confidence interval LB	UB
a_{55}	0.15335	0.01482	10.350	0.123	0.183
a_{56}	0.46408	0.26017	1.784	-0.060	0.988
a_{66}	0.49117	0.32315	1.520	-0.159	1.142
a_{67}	0.74722	0.22661	3.297	0.291	1.203
a_{77}	0.11479	0.28595	0.401	-0.461	0.690
a_{78}	0.55845	0.13339	4.186	0.290	0.827
a_{88}	0.35444	0.16506	2.147	0.221	0.686
a_{89}	0.85700	0.14098	6.079	0.573	1.141
a_{99}	0.07272	0.16184	0.449	-0.253	0.398
$a_{9,10}$	0.30289	0.16647	1.820	-0.032	0.638
$a_{10,10}$	0.58587	0.28358	2.066	0.015	1.157
$a_{10,11}$	0.67093	0.24710	2.715	0.174	1.168
$a_{11,11}$	0.41748	0.26288	1.588	-0.111	0.946

$t_{0.025} = 2.017$ $t_{0.05} = 1.680$

Table 7.8: Estimates of the Coefficients for the Unrestricted OLS for the Whole Country (Method I)

Variable	Coefficient Value	St.error	t-value	95% Confidence interval LB	UB
PEDUC1	0.08590	0.11703	0.734	-0.149	0.321
TEARN1	0.01624	0.01852	0.877	-0.021	0.053
PUTEA1	0.05573	0.15293	0.364	-0.252	0.364
PUCLASS1	-0.00289	0.02599	0.111	-0.055	0.049
BUS1	-0.05028	0.13236	0.380	-0.317	0.216
HELP1	-0.03961	0.04538	0.873	-0.131	0.0517
GDP1	0.07717	0.10246	0.753	-0.129	0.283
EARN1	-0.01032	0.03342	0.309	-0.078	0.057
POPLEV1	0.02684	0.11898	0.226	-0.213	0.266
PCAP2	-0.03835	0.07097	0.540	-0.181	0.104
TEARN2	-0.02842	0.02773	1.025	0.084	0.027
PUTEA2	-0.01825	0.11382	0.160	-0.247	0.211
HELP2	0.06671	0.08169	0.832	-0.095	0.228
GDP2	0.12364	0.08248	1.499	-0.042	0.289
LIFE2	-0.36471	0.26970	1.352	-0.907	0.178
UNEMP2	-0.05087	0.03273	1.554	-0.117	0.015
EARN2	0.05875	0.08201	0.716	-0.106	0.223
POPLEV2	0.13296	0.21303	0.624	-0.296	0.562

dF = 46 R = 0.99 R^2 = 0.96879 F = 314.86

Variables not entered in the Equation:

EDUC1 PCAP1 COST1 UNQUAL1 LIFE1 ILLIT1 UNEMP1 LFLEV1

EDUC2 PEDUC2 COST2 PUCLASS2 BUS2 UNQUAL2 ILLIT2 LFLEV2

$$\text{ILLIT} = 0.00091 - \underset{(9.07)}{0.94866}\ \text{POPLEV} \qquad R = .94941$$

(...)= t-values

The estimates of the coefficients $\underline{a}$, $\underline{\delta}_r$, $\underline{\delta}_p$ and the corresponding statistics have been used in FORTRAN programs STAT(VAR) and ERROR(CASE) (see Appendix E) to obtain the estimates of the different transition probabilities and their statistics.[3] These values are presented in Table 7.9 where the figures in brackets are the corresponding t-values.

The unrestricted OLS estimator does not guarantee that the non-negativity and row-sum conditions are satisfied. This can be seen in Table 7.9 where the results show that both conditions have been violated by this estimator. Furthermore, the t-values obtained are less than the critical value at the 5% significance level for most of the cases.

Although, in general, lower t-values are to be expected when constraints are imposed upon the model, more accurate estimates, however, must be obtained. Therefore, the next attempt is to apply the restricted OLS estimator to the extended model including all the seventeen explanatory variables and using stepwise regression.

7.4.2. The Restricted OLS Estimator

The straightforward use of the quadratic programming problem to achieve the restricted OLS estimates of the transition probabilities leads to a quite complicated objective function as more than twenty-four variables (seven elements in the $\underline{a}$ vector and the seventeen elements of the $\underline{\delta}_d$ vector) should be included in the model.

TABLE 7.9: Method I - Transition Probability Estimates for the Unrestricted OLS Method (Whole Country)

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.30 (4.36)	.62 (2.17)	.64 (1.88)	.91 (2.95)	.27 (.85)	.72 (2.32)	.50 (2.29)	1.02 (2.82)	.22 (.95)	.46 (1.00)	.73 (2.35)	.83 (1.76)	.57 (1.75)
1972	.28 (5.42)	.89 (2.50)	.62 (1.86)	1.18 (2.70)	.25 (.83)	.99 (2.02)	.48 (2.43)	1.29 (2.24)	.20 (1.04)	.73 (1.04)	.71 (2.20)	1.10 (1.53)	.55 (1.87)
1973	.25 (4.61)	.28 (.61)	.59 (1.77)	.57 (1.04)	.22 (.73)	.38 (.59)	.45 (2.29)	.68 (.93)	.17 (.82)	.12 (.14)	.68 (2.13)	.49 (.55)	.52 (1.75)
1974	.00 (.08)	.54 (1.59)	.34 (1.03)	.83 (2.19)	-.03 (.08)	.64 (1.64)	.20 (1.04)	.94 (2.06)	-.08 (.36)	.38 (.75)	.75 (1.35)	.75 (1.35)	.27 (.93)
1975	.07 (1.16)	.37 (1.18)	.41 (1.25)	.66 (2.06)	.04 (.15)	.47 (1.73)	.27 (1.33)	.77 (2.53)	-.01 (.03)	.21 (.67)	.50 (1.59)	.58 (1.55)	.34 (1.24)
1976	.04 (1.04)	.19 (.51)	.38 (1.17)	.48 (1.15)	.01 (.04)	.29 (.65)	.24 (1.31)	.59 (1.17)	-.04 (.21)	.03 (.06)	.47 (1.55)	.40 (.65)	.31 (1.12)
1977	.12 (2.83)	.46 (1.53)	.46 (1.44)	.75 (2.45)	.09 (.32)	.56 (1.94)	.32 (2.08)	.86 (2.76)	.04 (.24)	.30 (.89)	.55 (1.85)	.67 (1.47)	.39 (1.40)
1978	.15 (4.30)	.48 (1.75)	.49 (1.50)	.77 (2.81)	.12 (.41)	.58 (2.67)	.35 (1.94)	.88 (3.52)	.07 (.44)	.32 (1.18)	.58 (2.14)	.69 (1.85)	.42 (1.58)
1979	.16 (2.97)	.48 (1.57	.50 (1.51)	.77 (2.33)	.13 (.44)	.58 (1.82)	.36 (1.82)	.88 (2.42)	.08 (.39)	.32 (.79)	.59 (1.80)	.69 (1.37)	.43 (1.41)
1980	.16 (5.23)	.43 (1.12)	.50 (1.53)	.72 (1.52)	.13 (.44)	.53 (1.02)	.36 (2.02)	.83 (1.39)	.08 (.47)	.27 (.41)	.59 (2.07)	.64 (.82)	.43 (1.58)
1981	.11 (3.16)	.31 (.76)	.45 (1.38)	.60 (1.18)	.08 (.28)	.41 (.72)	.31 (1.72)	.71 (1.10)	.03 (.18)	.15 (.21)	.54 (1.83)	.52 (.62)	.38 (1.39)

Figures in parenthesis are t-values $t^{63}_{0.025} \approx 2.00$

Instead of forty seven exogenous variables, seventy-one exogenous variables should be used to generate the objective function of the programming problem. In this case the number of degrees of freedom is only sixteen. Furthermore, the number of constraint equations for the model exceeds the number of variables to be estimated. To overcome this problem, the constraint equations corresponding to the row-sum condition have been embodied in the model and the unrestricted OLS estimation procedure performed on the resultant model. This generates a new set of equations, corresponding to the global number of drop-outs observed per school year, which must be added to the previous equations.

The constraint equations to be embodied in the model, which impose the row-sum condition, can be written in the following form

$$p^t_{ii} + p^t_{i,i+1} + p^t_{i,s+1} = 1 \qquad i = 1, \dots ,s-1$$

$$t = 1, \dots ,k$$

$$p^t_{ss} + p^t_{s,s+1} = 1$$

that is,

$$\begin{aligned} p^t_{i,s+1} &= 1 - p^t_{ii} - p^t_{i,i+1} \\ p^t_{s,s+1} &= 1 - p^t_{ss} \end{aligned} \qquad (7.6)$$

Using equations (7.2), equations (7.6) can be rewritten in the form

$$p^t_{i,s+1} = 1 - a_{ii} - \underline{x}^{t'} \underline{\delta}_r - a_{i,i+1} - \underline{x}^{t'} \underline{\delta}_p$$

$$p^t_{s,s+1} = 1 - a_{ss} - \underline{x}^{t'} \underline{\delta}_r \qquad i=1, \ldots, k$$

$$t=1, \ldots, k$$

Thus, the last equation of the model given by equations (7.1) can be written as

$$n^t_{s+1} = \sum_{i=1}^{s} n_i^{t-1} - \sum_{i=1}^{s} a_{ii}\, n_i^{t-1} - \sum_{i=1}^{s-1} a_{i,i+1}\, n_i^{t-1} - \sum_{i=1}^{s} n_i^{t-1}\, \underline{x}^{t'}\, \underline{\delta}_r$$

$$-\sum_{i=1}^{s-1} n_i^{t-1}\, \underline{x}^{t'}\, \underline{\delta}_p + u^t_{s+1} \qquad t=1, \ldots, k$$

and the model previously defined by equations (7.3) has the new form

$$n^{*t}_1 = a_{11}\, n_1^{t-1} + n_1^{t-1}\, \underline{x}^{t'}\, \underline{\delta}_r + u^t_1 \qquad (7.7)$$

$$n^t_i = a_{i-1,i}\, n^{t-1}_{i-1} + a_{ii}\, n_i^{t-1} + (n_i^{t-1}\, \underline{x}^{t'}\ ,\ n^{t-1}_{i-1}\, \underline{x}^{t'}) \begin{bmatrix} \underline{\delta}_r \\ \underline{\delta}_p \end{bmatrix} + u^t_i$$

$$i=2, \ldots, s$$

$$n^{*t}_{s+1} = -\sum_{i=1}^{s} a_{ii}\, n_i^{t-1} - \sum_{i=1}^{s+1} a_{i,i+1}\, n_i^{t-1}$$

$$-\left(\sum_{i=1}^{s} n_i^{t-1}\, \underline{x}^{t'}\ ,\ \sum_{i=1}^{s-1} n_i^{t-1}\, \underline{x}^{t'}\right) \begin{bmatrix} \underline{\delta}_r \\ \underline{\delta}_p \end{bmatrix} + u^t_{s+1}$$

$$t=1. \ldots, k$$

where

$$n^{*t}_{s+1} = n^{t}_{s+1} - \sum_{i=1}^{s} n_i^{t-1}$$

Stacking over time and over variables, the model still can be described by equation (7.4)

$$\underline{n}^{*} = \underline{N}^{x}\,\underline{\delta} + \underline{u}$$

where $\underline{n}^{*}$ is the (88 x 1) vector with the form

$$\underline{n}^{*} = \begin{bmatrix} \underline{n}_1^{*} \\ \underline{n}_2 \\ \cdot \\ \cdot \\ \cdot \\ \underline{n}_s \\ \underline{n}_{s+1}^{*} \end{bmatrix}$$

and the matrix $\underline{N}^{x}$ is, in the same way, expressed by the two block matrices

$$\underline{N}^{x} = [\ \underline{N}^{o} \ \vdots\ \underline{n}^{t-1} \otimes \underline{x}^{t}\ ]$$

where

$$\underline{N}^{o} = \begin{bmatrix}
\underline{n}_1^{t-1} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \cdots & \underline{0} & \underline{0} \\
\underline{0} & \underline{n}_1^{t-1} & \underline{n}_2^{t-1} & \underline{0} & \underline{0} & & \cdot & \cdot \\
\cdot & \underline{0} & \underline{0} & \underline{n}_2^{t-1} & \underline{n}_3^{t-1} & & \cdot & \cdot \\
\cdot & \cdot & \cdot & \underline{0} & \underline{0} & & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & & \cdot & \cdot \\
\cdot & \cdot & \cdot & \cdot & \cdot & & \underline{0} & \underline{0} \\
\underline{0} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & & \underline{n}_{s-1}^{t-1} & \underline{n}_s^{t-1} \\
-\underline{n}_1^{t-1} & -\underline{n}_1^{t-1} & -\underline{n}_2^{t-1} & -\underline{n}_2^{t-1} & -\underline{n}_3^{t-1} & \cdots & -\underline{n}_{s-1}^{t-1} & -\underline{n}_s^{t-1}
\end{bmatrix}$$

and

$$\underline{n}^{t-1} \otimes \underline{x}^{t} = \begin{bmatrix} \underline{n}_1^{t-1} \otimes \underline{x}^t & \underline{0} \\ \underline{n}_2^{t-1} \otimes \underline{x}^t & \underline{n}_1^{t-1} \otimes \underline{x}^t \\ \underline{n}_3^{t-1} \otimes \underline{x}^t & \underline{n}_2^{t-1} \otimes \underline{x}^t \\ \vdots & \vdots \\ \underline{n}_s^{t-1} \otimes \underline{x}^t & \underline{n}_{s-1}^{t-1} \otimes \underline{x}^t \\ -\sum_{i=1}^{s} \underline{n}_i^{t-1} \otimes \underline{x}^t & -\sum_{i=1}^{s} \underline{n}_i^{t-1} \otimes \underline{x}^t \end{bmatrix}$$

Matrix $\underline{N}^x$ now has size (88 x 47), with $\underline{N}^o$ a submatrix of size (88 x 13) and $\underline{n}^{t-1} \otimes \underline{x}^t$ a submatrix of size (88 x 34).

It is, however, important to note that only the row-sum condition has been included in the model; the non-negativity condition has been left aside, as only a mathematical programming approach can include this kind of constraint. The job program JOBB with the FORTRAN program PROGRM(RESVAR) (see Appendix E), have been used to generate the new data file for the restricted OLS estimation procedure. The subprogram REGRESSION from the SPSS package of programs was applied to these new data, using the same stepwise method selected for the unrestricted OLS estimator. The results obtained are presented in Table 7.10 and Table 7.11. It is noticeable from this last table that of the five variables (COST, BUS, UNQUAL, ILLIT, LFLEV) excluded from the model, four of them (COST, UNQUAL, ILLIT, LFLEV) were also removed when performing the unrestricted OLS method.

Table 7.10: Estimate of the $\underline{a}$ Coefficients for the Restricted OLS for the Whole Country (Method I)

Coefficient	Value	St.error	t-value	95% Confidence interval LB	UB
a_{55}	0.15279	0.01271	12.024	0.127	0.178
a_{56}	0.47382	0.23291	2.034	0.007	0.940
a_{66}	0.47813	0.28852	1.657	-0.100	1.056
a_{67}	0.74959	0.20704	3.621	0.335	1.164
a_{77}	0.11076	0.26104	0.424	-0.411	0.633
a_{78}	0.56265	0.12125	4.640	0.320	0.805
a_{88}	0.34834	0.14993	2.323	0.048	0.648
a_{89}	0.85364	0.12842	6.647	0.596	1.110
a_{99}	0.07633	0.14812	0.515	-0.220	0.373
$a_{9,10}$	0.30005	0.15221	1.971	-0.005	0.604
$a_{10,10}$	0.59042	0.25868	2.282	0.072	1.108
$a_{10,11}$	0.65350	0.21909	2.983	0.214	1.092
$a_{11,11}$	0.43598	0.23438	1.861	-0.033	0.905

$t_{0.025} = 2.003$ $t_{0.05} = 1.673$

Table 7.11: Estimates of the δ Coefficients for the Restricted OLS for the Whole Country (Method I)

Variable	Coefficient Value	St.error	t-value	95% Confidence interval LB	UB
EDUC1	0.01482	0.01927	0.769	-0.024	0.053
PEDUC1	0.01569	0.05560	0.282	-0.095	0.127
PUCAP1	0.01531	0.04047	0.378	-0.065	0.096
PUCLASS1	-0.00622	0.02620	-0.253	-0.059	0.046
HELP1	-0.04179	0.05614	-0.744	-0.154	0.070
UNEMP1	0.01179	0.03595	0.329	-0.060	0.083
EARN1	0.00585	0.03796	0.154	-0.070	0.082
POPLEV1	0.05140	0.03315	1.551	-0.015	0.117
EDUC2	-0.03191	0.02372	1.345	-0.079	0.015
PCAP2	-0.03021	0.06750	0.447	-0.165	0.105
TEARN2	-0.01683	0.00880	1.914	-0.034	0.001
PUTEA2	-0.11189	0.10899	1.027	-0.330	0.106
PUCLASS2	-0.00723	0.02424	0.298	-0.056	0.041
HELP2	0.04866	0.07108	0.685	-0.094	0.191
GDP2	0.14971	0.07647	1.958	-0.003	0.303
LIFE2	-0.30590	0.06831	4.478	-0.141	-0.169
UNEMP2	-0.07495	0.03730	2.010	-0.150	-0.000
EARN2	0.04648	0.04339	1.071	-0.040	0.133

dF = 57 R = 0.99963 R^2 = 0.99887 F = 2473.34

Variables not entered in the Equation:

COST1 TEARN1 PUTEA1 BUS1 UNQUAL1 GDP1 LIFE1 ILLIT1

LFLEV1 PEDUC2 COST2 BUS2 UNQUAL2 ILLIT2 LFLEV2 POPLEV2

Using the new estimates obtained for the elements of vector $\underline{\delta}$ and using also their corresponding statistics, the FORTRAN programs STAT(VAR) and ERROR(CASE) were applied to produce the estimates of the transition probabilities and the associated statistics. These values are presented in Table 7.12. A comparison between the results obtained by the two estimation procedures performed shows that, although the t-values observed for the constant-terms estimates are higher when the constraint equations are included in the model, the same statement cannot be made for the t-values associated with the transition probabilities; some of the t-values obtained using the restricted OLS method are better than the t-values obtained using the unrestricted OLS method and vice versa. However, in both methods the t-values corresponding to the transition probabilities p_{77} and p_{99} are significantly low.

Some of the transition probabilities present estimates very far from their observed point estimates. In an extreme situation are p_{66} and $p_{10,10}$ for which the estimates obtained are much higher than the observed point estimates, using either the unrestricted or the restricted OLS method. For an understanding of the behaviour of the transition probability estimates, their patterns are compared with the patterns of the point estimates in Figures 7.1 to 7.13. It stands out from these graphs that, in general, the restricted OLS estimates have time-patterns more similar to the point estimate time-patterns than the unrestricted OLS estimates. In absolute terms, one can say that both estimation methods produce estimated values that do not represent reality; however, the time-patterns of the restricted OLS seem to reflect the time-patterns of the observed point estimates, for some of the transition probabilities.

TABLE 7.12: Method I - Transition Probability Estimates for the Restricted OLS Method (WHole Country)

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.16	.43	.48	.71	.12	.52	.35	.81	.08	.26	.60	.61	.44
	(2.82)	(1.91)	(1.64)	(3.28)	(.42)	(3.31)	(1.89)	(5.05)	(.43)	(1.26)	(2.04)	(2.36)	(1.60)
1972	.15	.47	.47	.75	.10	.56	.34	.85	.07	.30	.58	.65	.43
	(2.98)	(2.04)	(1.60)	(3.60)	(.38)	(3.75)	(1.96)	(5.55)	(.37)	(1.60)	(2.04)	(2.54)	(1.62)
1973	.12	.53	.45	.80	.08	.61	.32	.91	.05	.35	.56	.71	.41
	(2.86)	(2.29)	(1.58)	(3.89)	(.30)	(4.26)	(1.79)	(5.74)	(.26)	(1.90)	(1.95)	(2.90)	(1.53)
1974	.07	.65	.39	.92	.02	.74	.26	1.03	-.01	.47	.50	.83	.35
	(1.67)	(2.72)	(1.33)	(4.17)	(.09)	(5.93)	(1.53)	(6.66)	(.06)	(2.69)	(1.84)	(3.57)	(1.37)
1975	.14	.53	.47	.80	.10	.61	.34	.91	.06	.35	.58	.71	.42
	(3.44)	(2.14)	(1.59)	(3.44)	(.37)	(4.69)	(2.06)	(6.13)	(.37)	(1.67)	(2.07)	(3.64)	(1.67)
1976	.10	.59	.43	.86	.06	.68	.30	.97	.02	.41	.54	.77	.38
	(3.57)	(2.49)	(1.46)	(3.87)	(.22)	(5.53)	(1.92)	(6.86)	(.15)	(2.52)	(2.03)	(3.25)	(1.61)
1977	.15	.46	.48	.73	.11	.55	.35	.84	.08	.28	.59	.64	.44
	(5.06)	(1.88)	(1.65)	(3.61)	(.43)	(3.63)	(2.42)	(6.60)	(.49)	(1.68)	(2.19)	(2.68)	(1.83)
1978	.17	.35	.50	.63	.13	.44	.37	.73	.10	.18	.61	.53	.46
	(5.63)	(1.48)	(1.70)	(2.84)	(.49)	(3.07)	(2.28)	(4.54)	(.68)	(1.80)	(2.48)	(2.11)	(1.97)
1979	.19	.39	.52	.67	.15	.48	.39	.77	.12	.22	.63	.57	.48
	(5.71)	(1.67)	(1.77)	(3.20)	(.56)	(3.53)	(2.38)	(5.02)	(.72)	(1.26)	(2.29)	(2.41)	(1.90)
1980	.22	.40	.55	.68	.18	.49	.42	.78	.15	.23	.66	.58	.51
	(8.12)	(1.70)	(1.87)	(3.25)	(.68)	(3.69)	(2.53)	(5.29)	(.98)	(1.51)	(2.57)	(2.53)	(2.14)
1981	.20	.41	.53	.68	.16	.50	.40	.79	.13	.23	.64	.59	.49
	(6.81)	(1.76)	(1.81)	(3.30)	(.61)	(3.75)	(2.46)	(5.33)	(.76)	(1.26)	(2.35)	(2.41)	(2.01)

Figures in parenthesis are t-values

Fig. 7.1-7.13 Comparison between the Patterns of the OLS Estimates of the Transition Probabilities (Method I) and the Patterns of the Corresponding Observed Point Estimates

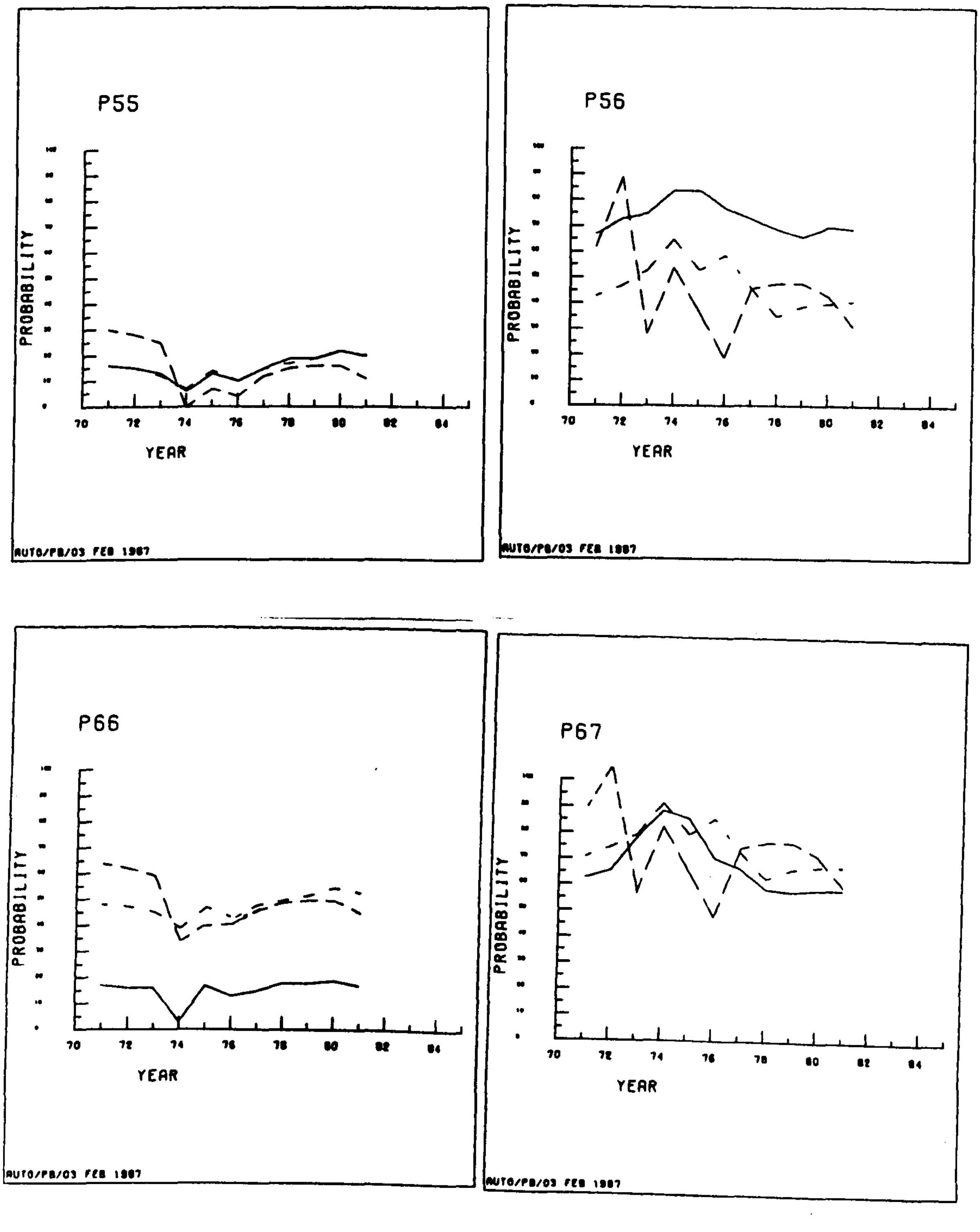


Fig. 7.1-7.13 Comparison between the Patterns of the OLS Estimates of the Transition Probabilities (Method I) and the Patterns of the Corresponding Observed Point Estimates (Continued)

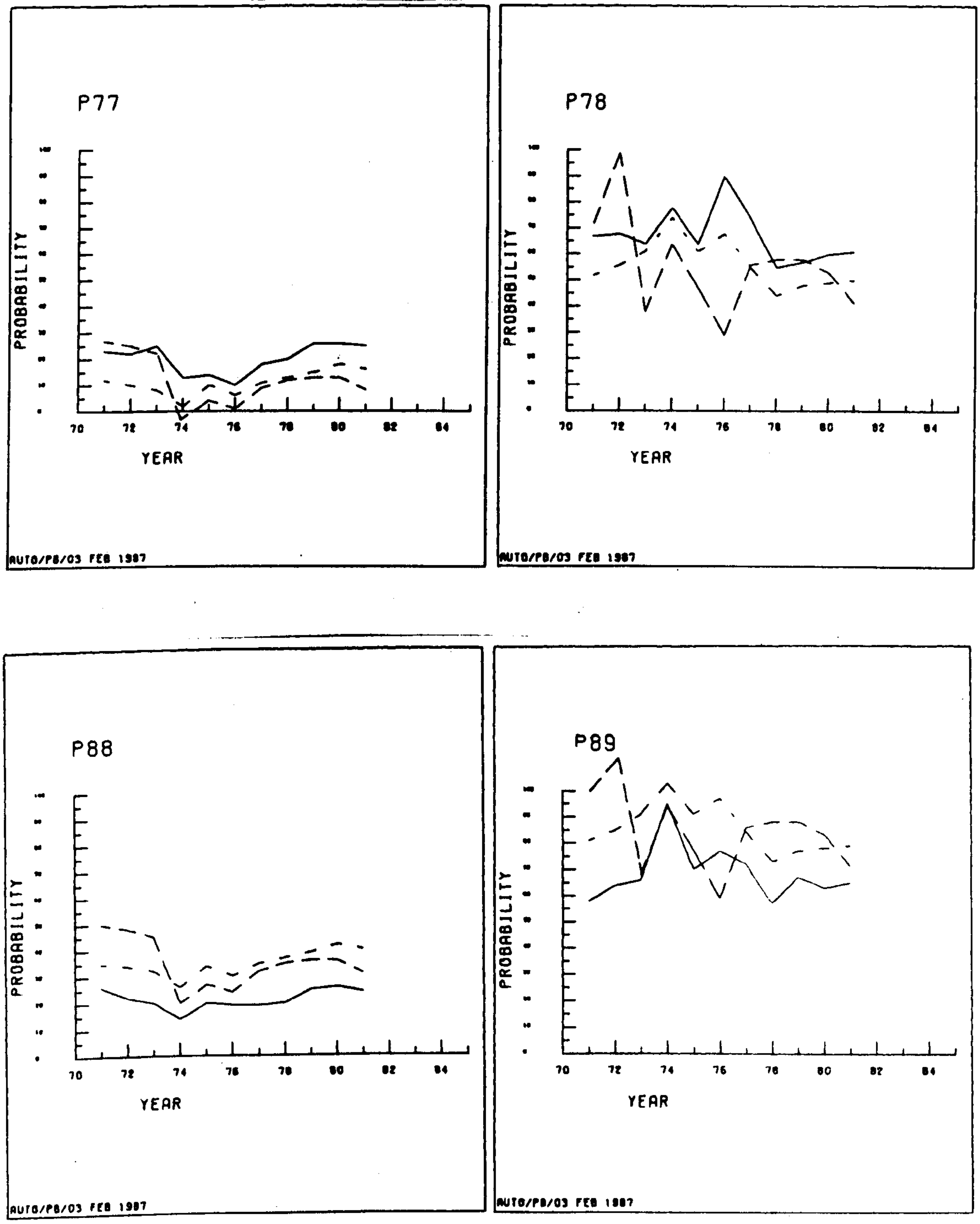
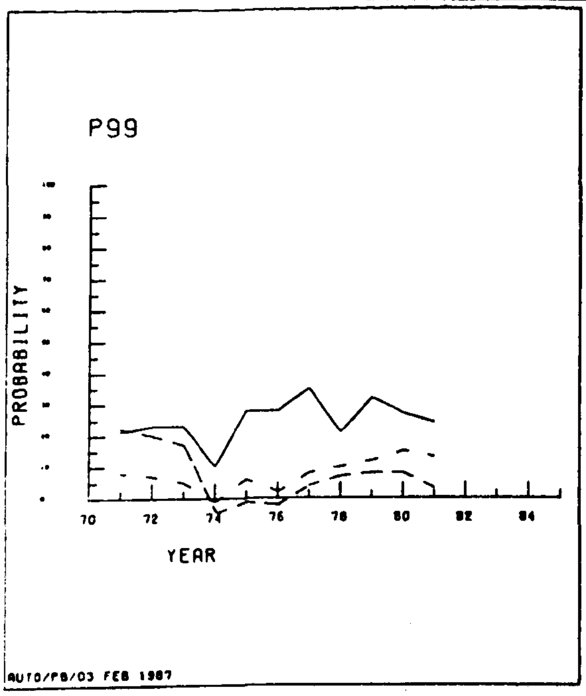
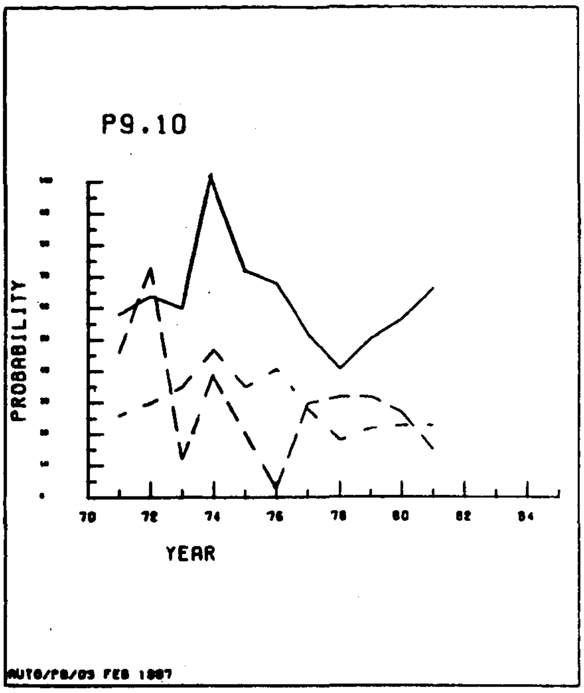
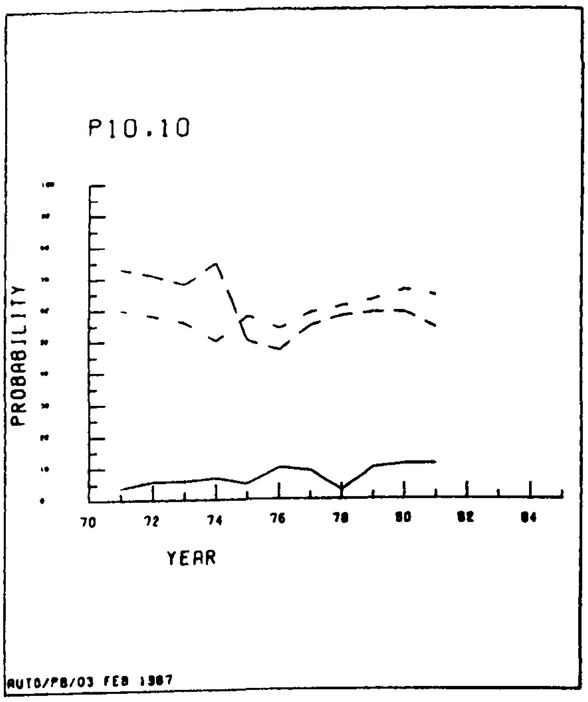


Fig. 7.1-7.13 **Comparison between the Patterns of the OLS Estimates of the Transition Probabilities (Method I) and the Patterns of the Corresponding Observed Point Estimates** (Continued)







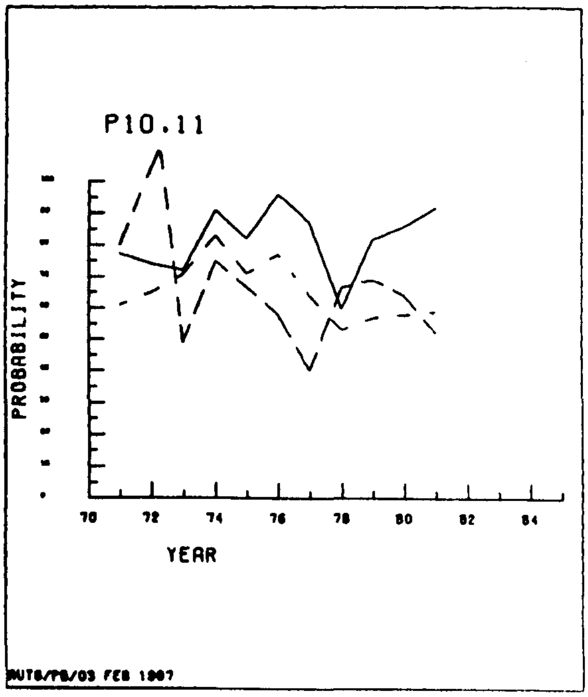
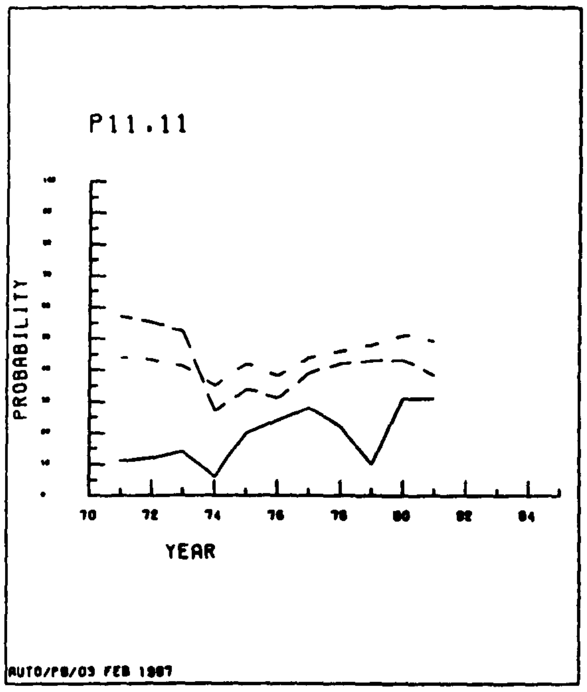


Fig. 7.1-7.13 **Comparison between the Patterns of the OLS Estimates of the Transition Probabilities (Method I) and the Patterns of the Corresponding Observed Point Estimates** (Continued)



——— Observed Point Estimates

— — — Unrestricted OLS Estimates

- - - - - Restricted OLS Estimates

These first results suggest that the study seems worthwhile and should be pursued in order to improve the analysis of the influence of the explanatory variables on the changes of the transition probabilities. The hypothesis of using explanatory variables in a Markov model to generate time varying transition probabilities seems to be a significant and important extension of the traditional Markov model. Alternative attempts should then be performed to estimate the parameters, trying, where possible, to overcome the problems of multicollinearity and heteroscedasticity, which occur in this kind of study. A regional disaggregation of the analysis, using district values of the explanatory variables could be useful, and is the subject of the next chapter.

7.5 Principal Components Analysis

7.5.1. Principal Components Analysis on the Explanatory Variables

The inclusion of all the explanatory variables separately in the the model leads to the acceptance of a certain degree of multicollinearity between the variables, even when performing the stepwise regression. This sub-section still retains all the explanatory variables and attempts instead to reduce their dimensionality and eliminate the interactions by using principal components analysis. This technique creates a smaller number of new variables which are linear combinations of the original ones. These components have the desirable statistical properties of being uncorrelated with each other and embodying a maximum amount of the variance of the original variables.

Subprogram FACTOR included in the SPSS package was used and applied to the data presented in Tables 7.4 and 7.5. For a better understanding and identification of the principal components to be obtained, FACTOR program has been applied separately to the supply side variables and to the demand side variables. The criterion used to extract the number of factors is known as the KAISER criterion [KAISER, 1960]. This approach advocates retaining as principal components only those factors that have eigenvalues greater than unity. This means that for a factor to be retained it must account for at least as much variance as does a single variable.

The starting point for principal components analysis is the correlation matrix. Useful initial information is obtained from a consideration of the largest correlation coefficients in the correlation matrix. As already noted in section 7.2, for statistical significance coefficients for a sample of eleven observations would need to be at least ± 0.7245 at the 1% level and ± 0.5961 at the 5% level.[4] However, it must be noted that with small samples, the correlation coefficients are quite unstable; the addition or omission of two or three observations can make a noticeable difference to the correlation value. Table 7.6 reveals that in the case of the demand side variables, only EARN has no significant values for the correlation coefficients, all the remaining variables being highly correlated. Similarly the supply side correlation matrix indicates that PUCLASS has no significant correlation with any other variable, and EDUC and TEARN have much more in common with each other than with the other variables.

The factor matrices obtained are presented in Table 7.13 and 7.14. The first ten and seven rows, respectively, contain the factor

TABLE 7.13: Factor Matrix and Rotated Factor Loadings for the Supply Side Explanatory Variables (Whole Country).

	Factor Matrix				Varimax Solution			Oblique Solution (Structure Matrix)		
	Common Factor Loading			Communality	Common Factor Loadings			Common Factor Loadings		
	I	II	III	h^2	I	II	III	I	II	III
EDUC	.06	.82	.30	.7606	- .08	.84	.22	- .00	.81	- .28
PEDUC	.95	.05	.06	.9129	.93	.20	.10	.95	.28	- .15
PCAP	- .76	- .17	.16	.6271	- .71	- .14	-.32	- .74	-.22	.36
COST	.86	- .08	.21	.7980	.87	.17	-.11	.87	.28	.05
TEARN	.30	.56	.48	.6335	.21	.78	-.04	.26	.78	- .04
PUTEA	- .99	.03	- .04	.9750	- .98	- .13	-.08	- .99	-.21	.13
PUCLASS	.25	.63	- .74	1.0017	.11	.13	-.99	.19	.21	-1.00
BUS	.93	- .33	.10	.9720	.97	- .09	+.14	.95	-.03	.11
HELP	.87	- .08	- .08	.7602	.86	- .00	.13	.87	-.08	- .17
UNQUAL	- .96	.09	.16	.9580	- .96	.04	-.20	- .96	-.05	.24
latent root	6.0115	1.7507	1.1025							
Percentage variance	60.1	17.5	11.0	88.6						

TABLE.7.14: Factor Matrix for the Demand Side Explanatory Variables (Whole Country)

	Common Factor Loading	Communality
	I	h^2
GDP	0.92	0.8417
LIFE	1.00	1.0050
ILLIT	-1.00	0.9978
UNEMP	0.87	0.7536
EARN	0.35	0.1257
LFLEV	0.99	0.9793
POPLEV	0.97	0.9491
latent root	5.738	
Percentage variance	82.0	90.2

loadings, which are the correlation coefficients between the observed variables and the composites. The row labeled 'latent root' (or eigenvalue) contains the sums of the squares of the factor loadings in each column. Because each of the factor loadings is a correlation and the square of a correlation is the amount of variance in one variable that is accounted for by the other variable, the latent root can be interpreted as the variance of the factor. The last row, which is labeled 'percentage of variance', is the eigenvalue for that factor (column) divided by the sum of the variances of all the variables. As these variables are in standardised form, each variable has a variance of 1.0; thus, the sum of these variances is equal to the number of observed variables included in the table.

The column labeled 'Communality' (h^2) contains the sum of squared factor loadings for each of the variables. These communalities give the amount of variance of each variable that is accounted for by the set of components. The 'uniqueness' of a variable can be seen also from Tables 7.13 and 7.14 as it is expressed as $1-h^2$; the uniqueness represents the sum of specific and error variance or unreliability. If the communality is too low, say in the region of 0.3 or less, giving a unique variance of 0.7 or more, it could well mean that the inclusion of that variable is unreliable, as the error variance could be making a major contribution. From the tables it can be seen that the lowest value for the communality (0.48) is for variable EARN, which justifies the non-elimination of any variable from the analysis.

The values 88.6 and 90.2 are, in each case, the proportion of variance in the total set of variables which is common variance. The high values observed for these common percentages of variance show that the principal components selected represent well the variations of the set of the original explanatory variables.

The main aim of principal components analysis is the identification of the components by examining the correlation with the original variables; that is, the factor loadings. For the purpose of specifying an acceptable level of significance, the loadings can be treated in a similar way to correlation coefficients. Loadings of at least +.59 and +.72 are recommended for the 5% and 1% levels, respectively. Because of the uncertainty about the magnitude of the error in factorial analysis, CHILD [1970, p.45] suggests it would be safer to adopt the 1% level as a criterion for significance of the loadings, especially in the case of small samples.

Two modifications of the general factor model were performed to the case of the supply side factors, to help the interpretation of the factors. These were transforming the location of the initial factors by a rotation procedure, either orthogonal (or Varimax) or oblique. The solutions can also be seen in Table 7.13.

The patterns of significant or non-significant loadings support the interpretation of the components.[5] The supply side determinants chosen in order to specify changes in the quality of schooling behaviour and the three principal factors can be broadly identified as follows:

Factor I - represents the economic characteristics of the supply for education as it is highly positively correlated with the costs of education, availability of scholarships and subsidies, and transport access offered to the students.

It is highly negatively correlated with the qualification and number of teachers. This component alone accounts for over two-thirds of the pooled variance.

Factor II - is representative of the teachers' motivation, as it has significant correlation coefficients with both teachers' salaries and education expenditures.

Factor III - can be identified as an inverse measure of school places offered to the students, being highly negatively correlated with the pupil-classroom ratio.

In the same way, the determinants of the students' demand for education is identified by one factor only. This factor represents a general well-being of the population, reflecting associated characteristics such as the unemployment, income and the educational levels of the population and the labour force.

The four principal components obtained and presented in Table 7.15 have been scaled to have zero mean and unit variance. The correlation coefficients computed between the supply side and the demand side principal components show that all the four new explanatory variables are independent of one another.

TABLE 7.15: The Principal Components (**Standardised Values**) of the Explanatory Variables - Whole Country.

YEAR	Supply Side I	Supply Side II	Supply Side III	Demand Side I
1971	-1.5720	0.7824	0.1797	-1.4172
1972	-1.0232	0.5396	0.6803	-1.1565
1973	-0.6547	-0.1464	-0.0882	-0.3629
1974	-0.7228	-0.2176	0.4986	-0.8708
1975	0.9706	-0.9148	-1.5743	-0.4014
1976	-0.7045	1.9526	1.9150	2.0059
1977	0.4569	-0.1811	-0.0566	0.3924
1978	1.0206	-1.3652	-0.8870	0.1283
1979	-0.0142	0.8695	0.7949	0.3132
1980	0.7179	-0.0055	-0.1770	0.2532
1981	1.5254	-1.3135	-1.2854	1.1158

Correlation coefficient

		Supply	factors	
Demand factor		I	II	III
	I	.46	.13	.17

7.5.2 The OLS Estimator with Principal Components

Using the four principal components as the explanatory variables of the extended model, new data files have been generated and new estimates for the $\underline{\delta}$ coefficients have been computed (Method II). Unrestricted and restricted OLS regressions have been performed and the results are presented in Table 7.16 and Table 7.17. In these Tables, S_{r1} to D_r represent the coefficients of the principal components (three supply side plus one demand side) affecting the repetition probabilities, and S_{p1} to D_p are the coefficients of the same principal components affecting the promotion probabilities. The estimates of the transition probabilities and corresponding t-values are shown in Table 7.18 and Table 7.19, and have been computed using the same procedure as for the previous estimations. It is clear from these tables, especially for the restricted OLS estimates, that the the non-negativity condition is strongly violated. In this case, very high estimates (greater than unity) for the transition probabilities and very low estimates (less than zero) for the repetition probability estimates have been found for the probabilities $\underline{\hat{p}}_{88}$ and onwards.

Figures 7.14 to 7.26[6] compare the patterns of the new OLS estimates with the observed point estimates. It can be seen from the

Table 7.16: Coefficients and Confidence Intervals for the Unrestricted OLS Estimation Procedure using Principal Components (Method II) (Whole Country)

Coefficient	Value	St.error	t-value	95% Confidence interval LB	UB
a_{55}	0.15455	0.01651	9.366	0.121	0.187
a_{56}	0.25323	0.23905	1.059	-0.225	0.732
a_{66}	0.75259	0.29664	2.537	0.158	1.346
a_{67}	0.59909	0.21463	2.752	0.163	1.035
a_{77}	0.30433	0.27348	1.112	-0.243	0.852
a_{78}	0.68896	0.13451	5.122	0.419	0.958
a_{88}	0.20093	0.16600	1.210	-0.131	0.533
a_{89}	0.93783	0.15002	6.251	0.637	1.238
a_{99}	-0.01862	0.17268	0.107	-0.364	0.327
$a_{9,10}$	0.13301	0.17840	1.745	-0.224	0.490
$a_{10,10}$	0.86396	0.30524	2.830	0.252	1.475
$a_{10,11}$	0.82327	0.27331	3.023	0.277	1.368
$a_{11,11}$	0.25985	0.29005	0.895	-0.321	0.840
Sr1	0.07466	0.04978	1.499	-0.025	0.174
Sr2	0.06873	0.04688	1.466	-0.025	0.162
Sr3	-0.03212	0.05406	0.594	-0.140	0.076
Dr1	-0.03265	0.03189	1.023	-0.096	0.031
Sp1	-0.1339	0.05182	2.584	-0.237	-0.0301
Sp2	-0.04812	0.04810	1.000	-0.144	0.048
Sp3	-0.01989	0.05592	0.355	-0.132	0.092
Dp1	0.05557	0.03264	1.703	-0.009	0.121

dF = 55 $t_{0.025} = 2.050$ $t_{0.05} = 1.673$

Table 7.17: Coefficients and Confidence Intervals for the Restricted OLS Estimation Procedure using Principal Components (Method II) (Whole Country)

Coefficient	Value	St.error	t-value	95% Confidence interval LB	UB
a_{55}	0.15387	0.01883	8.173	0.116	0.191
a_{56}	0.11840	0.23740	0.498	-0.355	0.592
a_{66}	0.91451	0.29480	3.102	0.326	1.503
a_{67}	0.56909	0.20937	2.718	0.151	0.987
a_{77}	0.33840	0.26267	1.288	-0.186	0.862
a_{78}	0.93142	0.12889	7.226	0.674	1.188
a_{88}	-0.10450	0.15854	0.659	-0.421	0.212
a_{89}	1.17272	0.15908	7.372	0.855	1.490
a_{99}	-0.30655	0.18349	1.671	-0.673	0.059
$a_{9,10}$	-0.23618	0.18593	1.270	-0.607	0.135
$a_{10,10}$	1.46720	0.32077	4.574	0.826	2.107
$a_{10,11}$	1.18397	0.29888	3.958	0.587	1.779
$a_{11,11}$	-0.14604	0.31927	0.457	-0.783	0.491
Sr1	0.08471	0.05662	1.496	-0.028	0.197
Sr2	0.07068	0.05348	1.321	-0.036	0.177
Sr3	-0.02202	0.06177	0.356	-0.145	0.101
Dr1	-0.03226	0.03622	0.890	-0.104	0.040
Sp1	-0.12041	0.06045	1.992	-0.241	0.000
Sp2	-0.09185	0.05668	1.620	-0.205	0.021
Sp3	-0.03414	0.06540	0.522	-0.096	0.164
Dp1	0.05455	0.03854	1.415	-0.022	0.131

dF = 66 $t_{0.050} = 1.668$ $t_{0.25} = 1.989$

TABLE 7.18: Method II - Transition Probability Estimates for the Unrestricted OLS Using Principal Components (Whole Country)

YEAR	p_{55}	p_{56}	p_{65}	p_{67}	p_{77}	p_{78}	p_{88}	p_{89}	p_{99}	$p_{9,10}$	$p_{10,10}$	$p_{10,11}$	$p_{11,11}$
1971	.13	.34	.73	.69	.28	.78	.18	1.03	-.04	.22	.84	.91	.24
	(2.44)	(1.47)	(2.40)	(2.98)	(.99)	(4.97)	(.88)	(6.20)	(.19)	(.93)	(2.46)	(3.27)	(.75)
1972	.13	.29	.73	.63	.28	.72	.18	.97	-.04	.17	.84	.86	.24
	(3.92)	(1.20)	(2.43)	(2.90)	(1.01)	(5.13)	(1.02)	(6.12)	(-.23)	(.90)	(2.68)	(3.05)	(.78)
1973	.11	.33	.71	.68	.26	.77	.16	1.01	-.06	.21	.82	.90	.22
	(3.12)	(1.41)	(2.36)	(3.01)	(.94)	(5.41)	(.87)	(6.53)	(.33)	(1.07)	(2.62)	(3.20)	(.72)
1974	.10	.30	.70	.65	.25	.74	.14	.99	-.08	.18	.81	.87	.20
	(2.42)	(1.25)	(2.30)	(2.94)	(.87)	(5.02)	(.82)	(5.76)	(.40)	(1.10)	(2.64)	(2.90)	(.66)
1975	.23	.18	.83	.52	.38	.61	.27	.86	.05	.06	.94	.75	.33
	(5.59)	(.73)	(2.77)	(2.32)	(1.36)	(4.07)	(1.47)	(5.31)	(.29)	(.26)	(2.90)	(2.82)	(1.13)
1976	.11	.33	.71	.67	.26	.76	.16	1.01	-.06	.21	.82	.90	.21
	(2.49)	(1.35)	(2.33)	(2.93)	(.92)	(4.98)	(.83)	(5.85)	(-.31)	(.96)	(2.50)	(2.95)	(.68)
1977	.17	.22	.76	.57	.32	.66	.21	.91	-.01	.10	.87	.79	.27
	(9.64)	(.92)	(2.57)	(2.62)	(1.15)	(4.80)	(1.29)	(5.81)	(-.05)	(.60)	(2.89)	(2.85)	(.93)
1978	.16	.21	.76	.55	.31	.64	.21	.89	-.01	.09	.87	.78	.27
	(5.37)	(.85)	(2.54)	(2.52)	(1.12)	(4.46)	(1.23)	(5.33)	(-.07)	(.55)	(2.94)	(2.68)	(.97)
1979	.18	.21	.78	.56	.33	.65	.22	.90	.00	.09	.89	.78	.28
	(5.86)	(.86)	(2.60)	(2.56)	(1.18)	(4.50)	(1.30)	(5.47)	(.02)	(.50)	(2.80)	(2.77)	(.97)
1980	.21	.17	.80	.52	.35	.61	.25	.86	.03	.05	.91	.74	.31
	(7.33)	(.71)	(2.70)	(2.38)	(1.29)	(4.23)	(1.48)	(5.28)	(.18)	(.29)	(2.93)	(2.70)	(1.06)
1981	.18	.20	.78	.55	.33	.64	.23	.88	.01	.08	.89	.77	.29
	(5.63)	(.82)	(2.61)	(2.44)	(1.20)	(4.32)	(1.31)	(5.35)	(.06)	(.43)	(2.92)	(2.71)	(.99)

Figures in parenthesis are t-values

TABLE 7.19: Method II - Transition Probability Estimates for the Restricted OLS using Principal Components (Whole Country)

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.12	.16	.88	.62	.30	.98	-.14	1.22	-.34	-.19	1.43	1.23	-.18
	(1.93)	(.70)	(2.90)	(2.64)	(1.07)	(5.76)	(.68)	(6.66)	(1.41)	(.76)	(3.93)	(4.05)	(-.49)
1972	.13	.15	.89	.60	.31	.97	.13	1.21	-.33	-.20	1.44	1.22	-.17
	(3.37)	(.64)	(2.98)	(2.85)	(1.16)	(2.01)	(.77)	(7.21)	(1.70)	(1.06)	(4.26)	(3.99)	(.51)
1973	.10	.19	.86	.64	.29	1.00	-.16	1.24	-.36	-.17	1.41	1.25	-.20
	(2.55)	(.80)	(2.90)	(2.91)	(1.06)	(6.95)	(.89)	(7.64)	(1.73)	(.83)	(4.25)	(4.11)	(.59)
1974	.09	.19	.86	.65	.28	1.01	-.16	1.25	-.37	-.16	1.41	1.26	-.21
	(2.04)	(.81)	(2.85)	(3.02)	(1.02)	(7.10)	(.96)	(7.09)	(1.88)	(.89)	(4.20)	(3.86)	(.61)
1975	.22	.01	.98	.46	.40	.82	-.04	1.06	-.24	-.34	1.53	1.07	-.08
	(4.70)	(.04)	(3.27)	(2.09)	(1.48)	(5.29)	(.21)	(5.78)	(1.14)	(1.51)	(4.62)	(3.66)	(.24)
1976	.13	.20	.89	.65	.31	1.01	-.13	1.25	-.34	-.16	1.44	1.26	-.17
	(2.51)	(.83)	(2.92)	(2.87)	(1.14)	(6.39)	(.71)	(6.56)	(1.50)	(.67)	(4.08)	(3.64)	(.51)
1977	.17	.10	.93	.55	.35	.91	-.09	1.15	-.29	-.26	1.48	1.16	-.13
	(8.42)	(.42)	(3.14)	(2.65)	(1.34)	(7.07)	(.58)	(6.94)	(1.64)	(1.40)	(4.66)	(3.78)	(.42)
1978	.16	.10	.92	.55	.34	.91	-.10	1.15	-.30	-.26	1.47	1.16	-.14
	(4.65)	(.40)	(3.08)	(2.61)	(1.29)	(6.72)	(.62)	(6.55)	(1.69)	(1.48)	(4.70)	(3.66)	(.45)
1979	.19	.08	.95	.54	.37	.90	-.07	1.14	-.27	-.27	1.50	1.15	-.11
	(3.44)	(.34)	(3.17)	(2.54)	(1.38)	(6.43)	(.43)	(6.32)	(1.43)	(1.32)	(4.49)	(3.62)	(.34)
1980	.21	.04	.97	.49	.39	.85	-.05	1.09	-.25	-.31	1.52	1.10	-.09
	(6.59)	(.16)	(3.25)	(2.34)	(1.48)	(6.14)	(.29)	(6.09)	(1.34)	(1.56)	(4.68)	(3.58)	(.28)
1981	.18	.07	.94	.52	.37	.89	-.08	1.13	-.28	-.28	1.50	1.14	-.12
	(4.92)	(.30)	(3.15)	(2.41)	(1.37)	(6.12)	(.44)	(6.26)	(1.44)	(1.41)	(4.75)	(3.59)	(.38)

Figures in parenthesis are t-value.

Fig. 7.14-7.26 Comparison between the OLS Estimates of the Transition Probabilities (using Principal Components) and the Observed Point Estimates

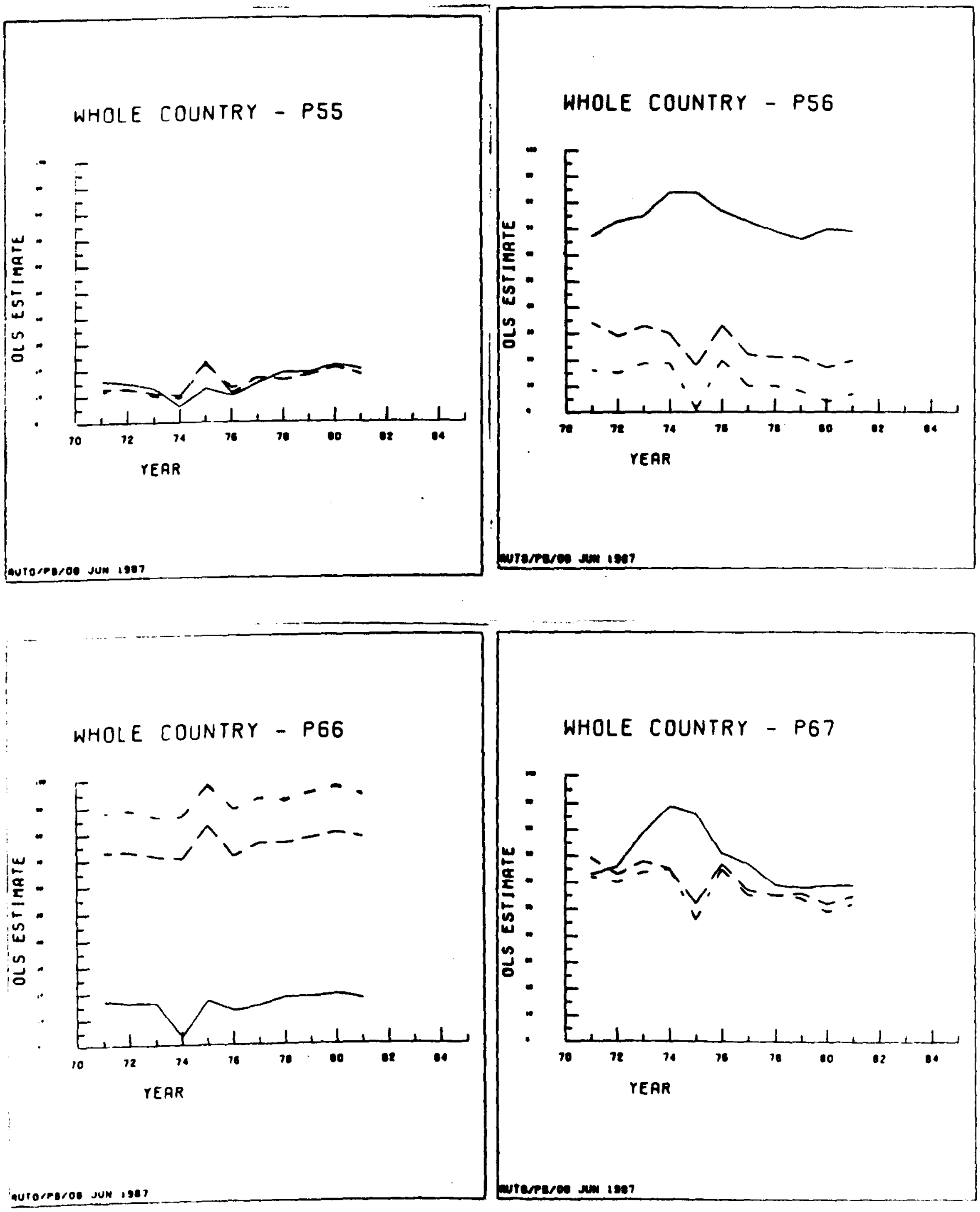


Fig. 7.14-7.26 Comparison between the OLS Estimates of the Transition Probabilities (using Principal Components) and the Observed Point Estimates (Continued)

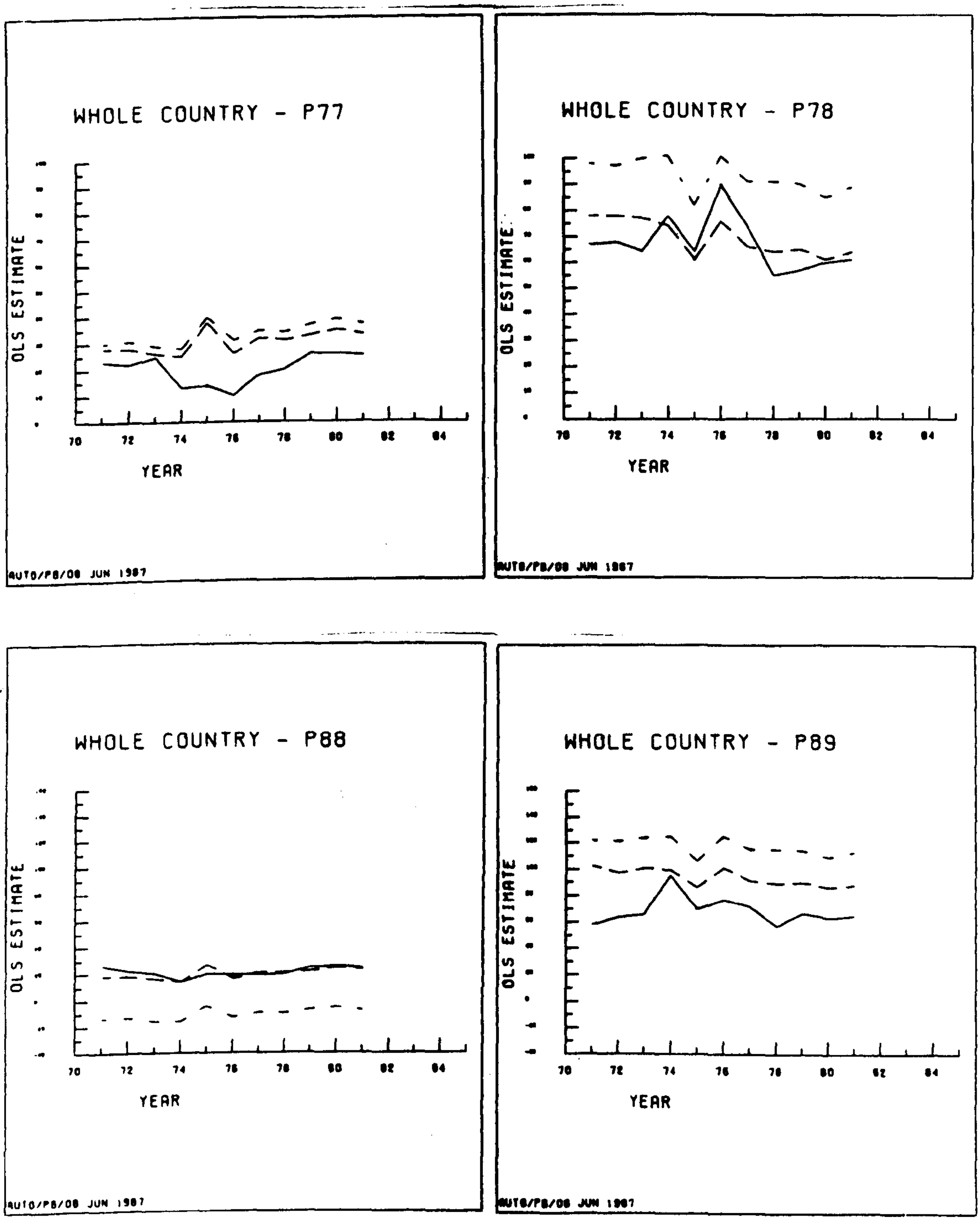
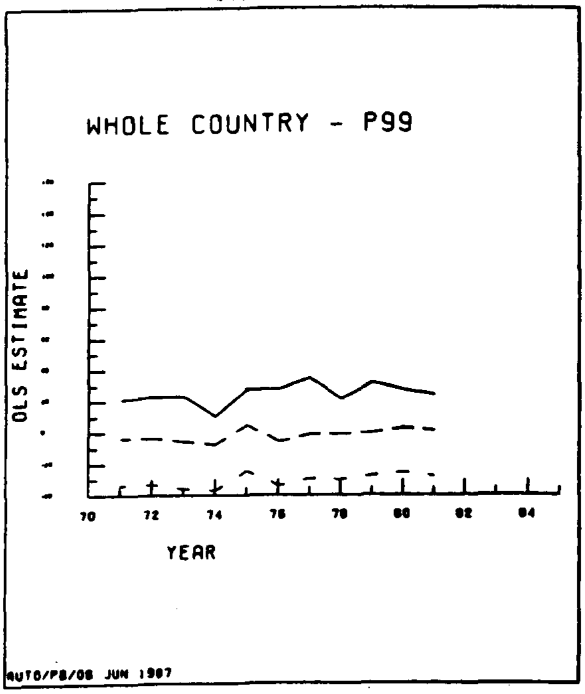
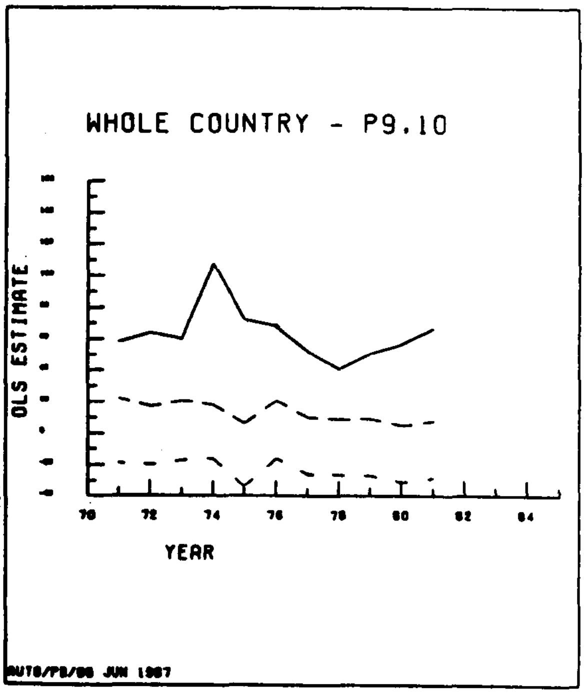
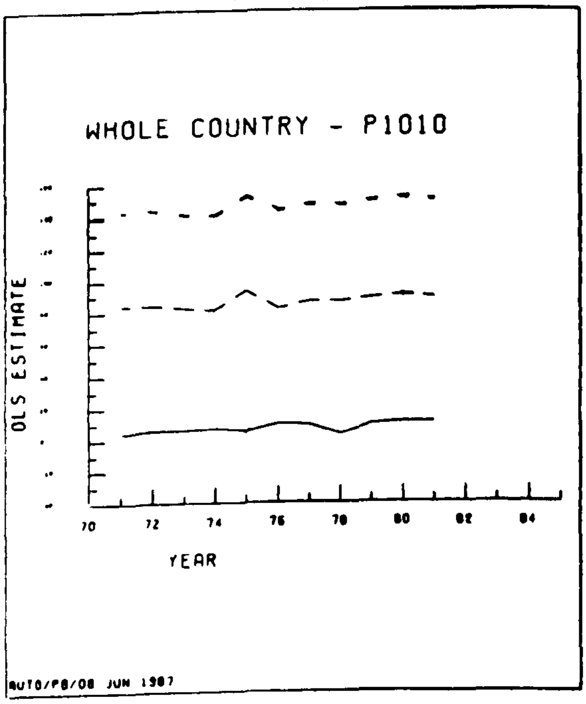


Fig. 7.14-7.26 Comparison between the OLS Estimates of the Transition Probabilities (using Principal Components) and the Observed Point Estimates (Continued)







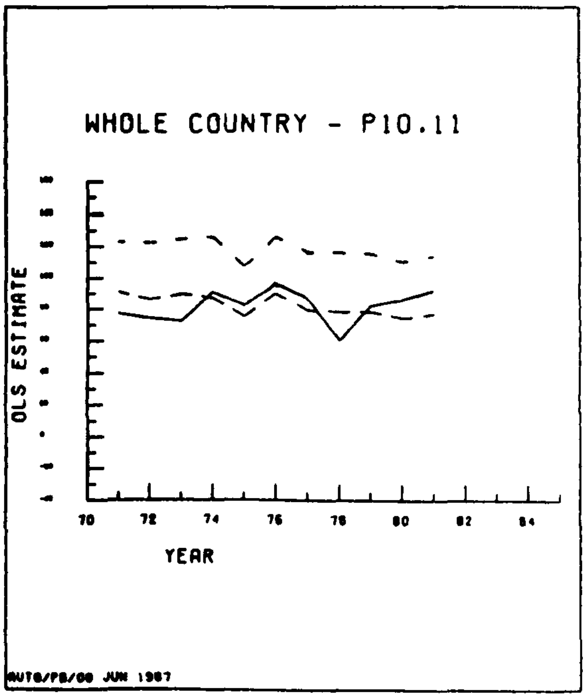
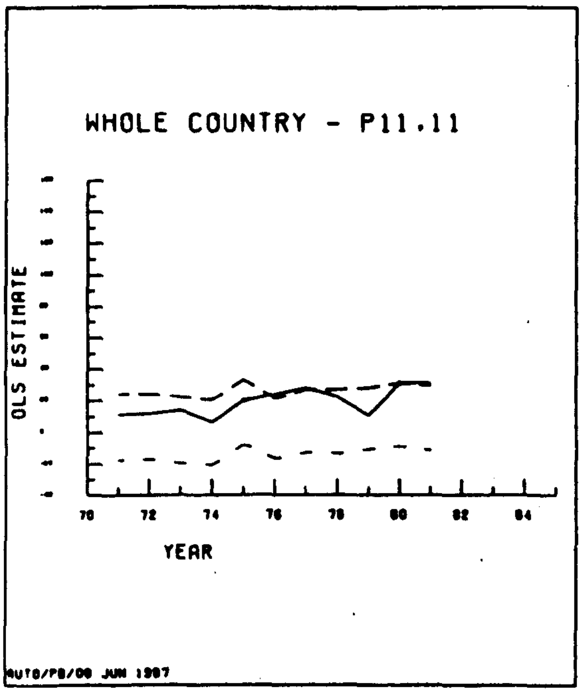


Fig. 7.14-7.26 Comparison between the OLS Estimates of the Transition Probabilities (using Principal Components) and the Observed Point Estimates (Continued)



——— Observed Point Estimates

— — — Unrestricted OLS Estimates

- - - - - Restricted OLS Estimates

graphs that the restricted OLS estimates are very far from compared to the observed point estimates. The differences observed when performing the unrestricted OLS method were increased with the inclusion of the row-sum condition. Furthermore, the patterns obtained using the principal components as the explanatory variables do not seem much more like the point estimate patterns than the patterns obtained when performing the stepwise regression on the set of original variables. Thus, although the use of principal components as explanatory variables has avoided the problem of multicollinearity, it has proved not to give more acceptable estimates for the present study.

A comparison between some statistics obtained performing the regression of the observed point transition probability estimates on each of the corresponding estimates is summarised in Table 7.20. The results presented in this table confirm the conclusions already taken from these estimation procedure approaches. The correlation coefficients are appreciably better when using the restricted OLS estimates from the stepwise regression (Method I). This validates the statement that although there is some discrepancies, in absolute values, between the observed point estimates and the restricted OLS stepwise regression estimates, these last estimates seem to represent most of the observed point estimates patterns satisfactorily. Also, the Durbin-Watson statistics shown in the table reveal that only these estimates give D-W values over the upper bound at the 5% significance level, which means that no significant serial correlation exists for the set of the transition probabilities. The same cannot be said for the other estimates obtained.

TABLE 7.20: Comparison Between the Estimates Obtained for the Different Estimation Procedures

Transition Probability	Observed Point Estimate		Method I - Stepwise Regression (Whole Country)							
			Unrestricted OLS				Restricted OLS			
	Mean Value	Stand Dev.	Mean Value	Stand Dev.	$R^{(a)}$	D-W	Mean Value	Stand Dev.	$R^{(a)}$	D-W
$P_{5,5}$	.1527	.0469	.1503	.0939	.42449	.42*	.1529	.0454	.98829	2.17
$P_{5,6}$	.7336	.0622	.4600	.1889	-.17442	.75*	.4728	.0901	.85931	2.45
$P_{6,6}$	.1536	.0441	.4903	.0939	.56014	1.57	.4784	.0454	.82517	2.42
$P_{6,7}$	.6873	.1123	.7500	.1889	-.14908	.49*	.7485	.0902	.87399	2.65
$P_{7,7}$	.2018	.0569	.1203	.0939	..69935	.50*	.1109	.0454	.75437	1.03
$P_{7,8}$	.6709	.1023	.5600	.1889	-.17098	1.73	.5615	.0902	.78880	1.61
$P_{8,8}$	.2136	.0367	.3503	.0939	.61116	.73*	.3485	.0454	.84519	1.23
$P_{8,9}$	.6855	.1050	.8600	.1889	-.14799	1.97	.8525	.0902	.85199	2.00
$P_{9,9}$	.2473	.0660	.0703	.0939	.07798	1.95	.0765	.0454	.48448	1.96
$P_{9,10}$	.6345	.1722	.3000	.1889	.01691	1.43	.2993	.0904	.83182	1.78
$P_{10,10}$	.0745	.0288	.5803	.0939	-.35506	1.58	.5905	.0455	.30738	1.93
$P_{10,11}$	.8173	.1044	.6700	.1889	-.38060	2.50	.6525	.0902	.48659	1.41
$P_{11,11}$	.1900	.0890	.4203	.0939	-.23588	1.44	.4360	.0452	.53540	1.27

(a) R is the correlation coefficient between the observed point estimates and the transition probability estimates

$D\text{-}W_{0.05}$ (1.00; 1.32)

TABLE 7.20: **Comparison Between the Estimates Obtained for the Different Estimation Procedures (Continued)**

Transition Probability	Method II - Regression on the Principal Components (Whole Country)							
	Unrestricted OLS				Restricted OLS			
	Mean Value	Stand Dev.	$R^{(a)}$	D-W	Mean Value	Stand Dev.	$R^{(a)}$	D-W
$P_{5,5}$	.1544	.0422	.64048	.74*	.1538	.0430	.64661	.72*
$P_{5,6}$	.2533	.0651	.12126	.70*	.1185	.0651	.14009	.72*
$P_{6,6}$	.7525	.0421	.61779	1.77	.9145	.0433	.60089	1.79
$P_{6,7}$	.5992	.0650	.29862	.76*	.5691	.0651	.28025	.79*
$P_{7,7}$	.3044	.0422	.26240	.77*	.3325	.0433	.21568	.78*
$P_{7,8}$	.6889	.0650	.58189	1.33	.9315	.0651	.61947	1.20
$P_{8,8}$	.2011	.0420	.46917	.85*	-.1045	.0433	.45046	.77*
$P_{8,9}$	.9378	.0650	.23851	1.62	1.1727	.0649	.37282	1.49
$P_{9,9}$	.0186	.0420	.52803	2.04	-.3065	.0433	.60951	2.13
$P_{9,10}$	.1331	.0650	.29337	1.48	-.2363	.0649	.33975	1.37
$P_{10,10}$	.8641	.0420	.19311	1.85	1.4670	.0430	.33618	2.14
$P_{10,11}$	.8234	.0650	.00090	1.83	1.1829	.0648	.03054	1.79
$P_{11,11}$	.2598	.0420	.53524	1.63	-.1460	.0430	.59574	1.82

(a) R is the correlation coefficient between the observed point estimates and the transition probability estimates

$D\text{-}W_{0.05}$ (1.00; 1.32)

Chapter 7

Footnotes

1. The unavailability of data concerning the rates of return to education for the different years covered by this study has prevented the inclusion of this indicator in the set of the original explanatory variables. As discussed in Chapter 2, MATILLA [1982] has shown that the rate of return is a strong and significant explanatory factor of the variation in the high school and college enrolment rates of young males.

2. These equations are similar to equations (4.1) presented in Chapter 4 for the basic model. The main difference between the two sets of equations is that in equations (7.1) the probabilities are not constant; they change over time.

3. To compute the statistics for the transition probabilities, equations (6.10) described in Chapter 6 have been used. This implies the use of the covariance matrix of the estimates coefficients.

4. CHILD [1970] includes in his book (p.95) the critical values for different sample sizes in a table titled 'Significance levels for Product Moment Correlation Coefficients'.

5. The use of principal components as a solution to multicollinearity has been widely questioned [see GLAUBER and FARRAR, 1967; pp.92-107]. The reasons for this is that the

method of principal components uses less of the information contained in the sample than does the normal OLS method applied to all explanatory variables. A corrective solution could involve the use of more information, which means the increase of the size of the sample (an impractical procedure in this study). However, if it is possible to attribute a clear meaning to the principal components, this loss of information is compensated for by the "meaningfulness" of the parameters of the model. This reduction is suggested when the number of variables is too large compared to the size of the sample.

6. Note that in these graphs the scale on the y axis is different for the probabilities p_{88} and onwards. Instead of the normal scale [0,100], the [-40,160] scale was used to include the OLS estimates corresponding to these probabilities.

CHAPTER 8

THE ESTIMATION OF THE PARAMETERS OF THE EXTENDED MODEL USING REGIONAL DATA

8.1. Introduction

Regional data are used in this chapter with the aim of improving the results of the application of the extended Markov model to the Portuguese educational system. The data corresponding to the different districts have been stacked and the data have been treated as a single set. Stepwise regression and pooled cross-section time-series analysis have been applied to produce the unrestricted and restricted OLS estimates of the different transition probabilities over time.

Principal components analysis was applied to each district individually. Using these principal components as the explanatory variables of the model, unrestricted and restricted OLS estimation procedures were performed on the new sets of stacked data, and new estimates of the time paths of the transition probabilities have been obtained.

Subsequently, an analysis of the results obtained and a comparison between the different methods is made. A summary description of the programs used in this chapter and displayed in Appendix F is presented in Appendix G.

8.2. The Explanatory Variables

Some differences have arisen while gathering the data concerning the explanatory variables at district level. The unavailability of data for some of the variables for all the years of this study, has led to the use of interpolation to establish estimates of the missing values that have occurred. Furthermore, the unavailability of data by district for variables LFLEV and POPLEV has led to the use of the whole country values for each district. Some variables, such as GDP, EDUC and TEARN have, by their nature, the same values for each district, and are equal to the whole country values. In a system with equality of opportunities and equal income distribution, variables PEDUC, PCAP and COST should behave in the same way as the above variables, and should have equal values for each district. Although this does not seem to be the situation for most countries, including Portugal, there were no available disaggregated data to make possible the calculation of these indicators. This has forced the assumption of equal values for variables PEDUC, PCAP and COST for all districts.

It can be seen, then, that only nine of the seventeen explanatory variables selected display different values by district, five variables on the supply-side and four variables on the demand-side of the education system. Therefore, although seventeen variables have been selected to study the changes over time of the transition probabilities, only nine of these variables can be considered potentially to explain some of the regional disparities observed in the behaviour over time of these transition probabilities. The remaining eight variables included in the analysis can explain changes in the transition probabilities that, by their nature, are national changes and so common to all districts equally. The observed data for these nine variables (PUTEA, PUCLASS, BUS, HELP UNQUAL, LIFE, ILLIT, UNEMP, EARN) are presented in Appendix F, Tables F.1-F.9; the corresponding standardised values are also presented in the same Appendix, Tables F.10-F.18.

From the tables some features are noticeable. The pupil-teacher ratio (PUTEA) has decreased rapidly during the 1970s, with a decrease also in the differences observed between districts. The pupil-classroom ratio (PUCLASS) presents, on the contrary, big increases and disparities between districts. It must be noted, however, that the big expansion of school enrolment, not followed by a growth in the capacity of the school system, has led to the use of the same physical space by two groups of students, one attending the school during the morning and the other attending school in the afternoon. The existence of this double (sometimes triple) use of classrooms by some schools (mainly in the more populated areas, which include, therefore, the big cities) is the reason why high values and big increases are observed in the pupil-classroom ratios.

The facilities of school bussing (BUS) and the number of scholarships (HELP) offered to the students show that there has been a huge increase in these variables since 1973, when these benefits started to be offered to students by the governmental institutions. It can be seen too from the tables that the percentage of students who use these facilities is higher in the districts that include areas with rural characteristics.

The percentage of non-qualified teachers (UNQUAL) is noticeably low in Lisboa, Porto and Coimbra (the three districts corresponding to the main cities of the country) immediately followed by their satellite districts. The more rural and interior districts are those which show higher percentages of non-qualified teachers. This is a consequence of the centralized policy of teachers' allocation that takes place every year at national level. Only those teachers with both the "academic" and "professional" qualifications can be given a permanent post on the staff of a school, and only these teachers have security of tenure; all the others have contracts for two years, at the end of which they can be displaced by anyone with a full qualification.

The district disparities observed in the early 1970s for the variable life expectation (LIFE) have been greatly reduced during the decade; this is probably attributable to the policy of decentralisation of medical care and of medical facilities. Finally, the illiteracy rate (ILLIT) has decreased in all districts during the period of analysis. This decrease can be attributed to the introduction of compulsory education for girls in the 1960s, together with a policy of encouragement of students to attend school. It must be noted that an intensification of recurrent education also took place in the second half of the 1970s.

8.3. The OLS Estimator

Data matrices have been generated for each district using the same process undertaken in the previous chapter for the aggregated data. Instead of performing the OLS estimation procedure for each district individually, as it was applied when analysing the efficacy of the basic Markov model, for this chapter the data matrices have been stacked over districts and the unrestricted and restricted OLS stepwise regressions have been performed on this new set of data

(Method III). This stacking of the data, not only over time and over grade but also over districts, has the aim of improving the reliability of the estimators, as the number of degrees of freedom increases greatly and the degree of multicollineanity between the independent variables should be appreciably reduced.

It has been assumed, however, that the explanatory variables have the same effect on the transition probabilities, the estimated values of the coefficients $\underline{a}$, $\underline{\delta}_m$, δ_p and $\underline{\delta}_d$ being then equal for each district. This means that only the district differences on the selected explanatory variables affect the changes of the problems of each type of transition probability. Thus, matrix $\underline{N}^x$ has now size (1386 x 47) for the unrestricted OLS estimation procedure and has size (1584 x 47) for the restricted OLS estimation procedure.

Stepwise regressions were performed in both cases, using SHAZAM, version 4.3, available at the University of Manchester Regional Computer Centre (see Appendix F, program REGSHAX for the unrestricted OLS regression; a similar program has been used to perform the restricted OLS regression). The results of these estimation procedures are presented in Table 8.1 and the corresponding transition probability estimates for the whole country are presented in Tables 8.2 and 8.3.

Although the estimates of the transition probabilities seem to be closer to the observed point estimates than the estimates obtained in the previous chapter, the two regressions used show that the estimates for the transition probabilities corresponding to terminal or first grades of levels of education ($\hat{p}_{67}$, $\hat{p}_{77}$, $\hat{p}_{9,10}$ and $\hat{p}_{10,10}$) are very low for promotion probabilities and very high for repetition

TABLE 8.1: Estimated Values of the $\underline{a}$ and δ Coefficients, using stacked District Data (Method III)

Unrestricted OLS				Restricted OLS			
Coefficient	Estimated Value	St. Error	t-value	Coefficient	Estimated Value	St. Error	t-value
a_{55}	.16	.0066	24.02	a_{55}	.14	.0082	17.73
a_{56}	.69	.0692	9.97	a_{56}	.73	.0711	10.22
a_{66}	.30	.0849	3.57	a_{66}	.24	.0874	2.73
a_{67}	.35	.0471	7.49	a_{67}	.20	.0483	4.15
a_{77}	.70	.0546	12.85	a_{77}	.86	.0560	15.33
a_{78}	.69	.0529	13.09	a_{78}	.79	.0629	12.53
a_{88}	.26	.0651	3.99	a_{88}	.12	.0776	1.51
a_{89}	.82	.0531	15.44	a_{89}	.72	.0610	11.74
a_{99}	.13	.0617	2.17	a_{99}	.23	.0709	3.26
$a_{9,10}$	.33	.0595	5.61	$a_{9,10}$	.44	.0647	6.84
$a_{10,10}$	.70	.0958	7.30	$a_{10,10}$	.49	.0143	4.65
$a_{10,11}$	.73	.0910	8.04	$a_{10,11}$	.52	.1075	4.88
$a_{11,11}$	.36	.0955	3.72	$a_{11,11}$	.53	.1126	4.75
EDUC 1	.0154	.0052	2.98	PEDUC 1	.0637	.0105	6.08
PUCLASS 1	-.0250	.0064	3.94	COST 1	.0375	.0075	5.00
GDP 1	.0354	.0063	5.62	PUTEA 1	-.0275	.0063	4.37
PCAP 2	-.0269	.0098	2.75	PUCLASS 1	-.0207	.0040	5.20
TEARN 2	-.0326	.0053	6.20	BUS 1	-.0292	.0083	3.51
PUTEA 2	-.0273	.0013	2.12	HELP 1	-.0306	.0110	2.77
HELP 2	.0352	.0081	4.34	GDP 1	.1685	.0157	10.74
GDP 2	.0597	.0181	3.30	ILLIT 1	-.1340	.0282	4.76
UNEMP 2	-.0726	.0091	7.94	UNEMP 1	-.0504	.0134	3.76
EARN 2	.0387	.0079	4.93	LFLEV 1	-.2701	.0249	10.84
LFLEV 2	-.1070	.0197	5.44	TEARN 1	-.0328	.0031	10.44
				HELP 2	.0408	.0118	3.47
				UNQUAL 2	.0477	.0090	5.30
				UNEMP 2	-.0389	.0122	3.20
				EARN 2	.0395	.0046	8.64

R^2 = .9660 | R^2 = .9932

$\bar{R}^2$ = .9655 | $\bar{R}^2$ = .9931

df = 1361 | df = 1555

$t_{0.025}$ = 1.96

TABLE 8.2: Method III - Transition Probability Estimates using Stepwise Regression Applied to all Stacked Districts with Explanatory Variables (Unrestricted OLS)-Whole Country.

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.18	.66	.32	.32	.72	.66	.28	.79	.15	.30	.72	.70	.38
	(9.16)	(9.59)	(3.44)	(6.06)	(11.08)	(11.28)	(3.73)	(13.01)	(1.98)	(4.53)	(6.89)	(7.12)	(3.61)
1972	.12	.68	.26	.34	.66	.68	.22	.81	.09	.32	.66	.72	.32
	(11.09)	(9.79)	(3.03)	(6.77)	(11.87)	(11.38)	(3.29)	(14.77)	(1.35)	(4.58)	(6.68)	(7.61)	(3.25)
1973	.16	.72	.30	.38	.70	.72	.26	.85	.13	.36	.70	.76	.36
	(18.61)	(10.72)	(3.51)	(7.70)	(12.66)	(13.41)	(3.99)	(14.38)	(1.97)	(5.47)	(7.22)	(7.73)	(3.71)
1974	.09	.84	.23	.50	.63	.84	.19	.97	.06	.48	.63	.88	.29
	(7.29)	(12.01)	(2.71)	(10.35)	(11.17)	(15.14)	(2.78)	(18.47)	(1.01)	(6.98)	(6.30)	(9.56)	(2.96)
1975	.15	.73	.29	.39	.69	.73	.25	.86	.12	.37	.69	.77	.35
	(9.47)	(10.19)	(3.44)	(6.89)	(17.75)	(13.91)	(3.76)	(14.44)	(1.61)	(4.61)	(6.77)	(8.86)	(3.44)
1976	.12	.72	.26	.38	.66	.72	.22	.85	.09	.36	.66	.76	.32
	(12.23)	(10.15)	(3.00)	(7.49)	(11.87)	(13.21)	(3.26)	(15.09)	(1.44)	(5.77)	(6.75)	(8.11)	(3.30)
1977	.15	.72	.29	.38	.69	.72	.25	.85	.12	.36	.69	.76	.35
	(20.80)	(9.84)	(3.42)	(7.96)	(12.63)	(11.20)	((3.85)	(15.93)	(1.94)	(5.83)	(7.18)	(7.70)	(3.66)
1978	.16	.60	.30	.26	.70	.60	.26	.73	.13	.24	.70	.64	.76
	(21.06)	(8.45)	(3.45)	(5.26)	(12.63)	(10.32)	(3.82)	(11.68)	(2.10)	(4.74)	(7.22)	(6.51)	(3.71)
1979	.19	.64	.33	.30	.73	.64	.29	.77	.16	.28	.73	.68	.39
	(23.95)	(8.95)	(3.90)	(6.16)	(13.12)	(11.27)	(4.33)	(12.43)	(2.61)	(4.40)	(7.57)	(7.18)	(4.06)
1980	.23	.63	.37	.29	.77	.63	.33	.76	.20	.27	.77	.67	.43
	(19.35)	(9.12)	(4.39)	(6.04)	(13.28)	(11.54)	(4.85)	(12.12)	(2.89)	(4.51)	(7.81)	(6.86)	(4.35)
1981	.22	.65	.36	.31	.76	.65	.32	.78	.19	.29	.76	.69	.42
	(19.18)	(9.30)	(4.16)	(6.25)	(13.05)	(11.30)	(4.55)	(13.02)	(2.94)	(4.28)	(7.71)	(6.93)	(4.25)

Figures in parenthesis are t-values

TABLE 8.3: Method III - Transition Probability Estimates Using Stepwise Regression Applied to All Stacked Districts with Explanatory Variables (Restricted OLS) - Whole Country

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.13	.76	.23	.23	.85	.82	.10	.75	.22	.47	.48	.55	.52
	(6.06)	(10.40)	(2.46)	(3.89)	(13.07)	(10.88)	(1.16)	(11.38)	(2.67)	(5.74)	(4.18)	(4.68)	(4.29)
1972	.10	.72	,20	.19	.82	.78	.07	.71	.19	.43	.45	.51	.49
	(4.73)	(9.86)	(2.27)	(3.33)	(12.34)	(10.69)	(.83)	(10.74)	(2.29)	(5.68)	(4.13)	(4.32)	(3.81)
1973	.11	.77	.21	.24	.83	.83	.08	.76	.20	.48	.46	.56	.50
	(6.17)	(10.71)	(2.31)	(4.44)	(13.23)	(12.11)	(.94)	(11.74)	(2.53)	(5.87)	(3.80)	(4.68)	(3.84)
1974	.06	.86	.16	.33	.78	.92	.03	.85	.15	.57	.41	.65	.45
	(4.63)	(11.90)	(1.83)	(6.64)	(13.70)	(14.25)	(.37)	(13.49)	(2.02)	(8.15)	(3.71)	(5.94)	(3.90)
1975	.14	.73	.24	.20	.86	.79	.11	.72	.23	.44	.49	.52	.53
	(10.65)	(10.26)	(2.87)	(3.54)	(14.28)	(14.95)	(1.54)	(9.83)	(2.70)	(5.72)	(4.74)	(5.51)	(4.48)
1976	.10	.76	.20	.23	.82	.82	.07	.75	.19	.47	.45	.55	.49
	(7.75)	(10.65)	(2.22)	(4.64)	(14.47)	(13.30)	(.84)	(11.28)	(2.36)	(6.52)	(3.96)	(5.19)	(4.55)
1977	.15	.74	.25	.21	.87	.80	.12	.73	.24	.45	.50	.53	.54
	(15.16)	(10.14)	(2.96)	(4.50)	(15.28)	(11.52)	(1.77)	(11.92)	(3.40)	(7.67)	(4.85)	(4.62)	(4.94)
1978	.13	.66	.23	.13	.85	.72	.10	.65	.22	.37	.48	.45	.52
	(10.03)	(9.06)	(2.56)	(2.64)	(14.93)	(10.49)	(1.25)	(9,85)	(3.41)	(6.14)	(4.44)	(4.05)	(4.88)
1979	.23	.64	.33	.11	.95	.70	.20	.63	.32	.35	.58	.43	.62
	(16.99)	(8.74)	(3.64)	(2.16)	(15.80)	(10.15)	(2.32)	(9.40)	(4.61)	(5.70)	(5.51)	(3.31)	(5.27)
1980	.23	.68	.33	.15	.95	.74	.20	.67	.32	.39	.58	.47	.62
	(21.12)	(9.38)	(3.70)	(3.04)	(15.80)	(11.49)	(2.48)	(10.08)	(4.34)	(6.07)	(5.80)	(4.20)	(5.24)
1981	.15	.73	.25	.20	.87	.79	.12	.72	.24	.44	.50	.52	.54
	(11.11)	(10.15)	(2.76)	(4.09)	(14.72)	(12.13)	(1.49)	(10.64)	(3.29)	(6.04)	(4.75)	(4.61)	(4.97)

Figures in parenthesis are t-values

probabilities. Similar results have also been observed for most of the estimation procedures previously attempted.

As mentioned in Chapter 6, the non-negativity condition is not imposed when the extended model is applied to the present study. Nevertheless, Tables 8.2 and 8.3 show that although the restricted OLS repetition probability estimates and the restricted OLS promotion probability estimates are positive, the imposed row-sum condition produces negative drop-out probability estimates for the seventh and the tenth grades ($\hat{p}_{7d}$ and $\hat{p}_{10d}$). This is a consequence of the corresponding very high observed repetition probability estimates. As the unrestricted OLS estimates should also satisfy the row-sum condition, it is possible to use this condition to obtain the unrestricted OLS drop-out probability estimates. It results in only one more grade having negative drop-out probability estimates ($\hat{p}_{9,d}$), this grade being the terminal grade of the secondary unified general course, affected by final examinations.

It is noted, however, from these tables that, with the increase in the number of observations of the system, the process of stacking the data over districts has given significant (5%) t-values for all coefficients except $\hat{a}_{88}$ for the restricted OLS estimation procedure. For the transition probabilities, exceptions are observed in the repetition probability estimate $\hat{p}_{99}$ for the unrestricted OLS estimation procedure and in the repetition probability estimate $\hat{p}_{88}$ for the restricted OLS estimation procedure. Nevertheless, it must be noted that only four out of eleven t-values are non-significant in the case of $\hat{p}_{99}$ and all t-values presented in Table 8.1 are significant.

Figures 8.1-8.13 compare the time-patterns of these transition probability estimates (Method III) with the corresponding observed point estimates. The graphs show that contrary to the results obtained in the previous chapter for the Method I and Method II, the unrestricted OLS estimation procedure using stacked district data gives time-patterns more similar to the observed point estimates time-patterns than the restricted OLS estimation procedure. Table 8.4 presents some of the statistics obtained on performing the regression of the transition probabilities observed point estimates on the corresponding stepwise regression estimates using the stacked district data. The table confirms the conclusions taken from the graphs, with the correlation coefficients using the unrestricted OLS estimate in general being higher than the correlation coefficients using the restricted OLS estimates. The Durbin-Watson statistic shows the existence of positive serial correlation for the repetition probability $\hat{\underline{p}}_{77}$; however, the remaining probabilities exhibit no serial correlation.

A comparison between Table 8.4 and Table 7.20 shows that with the exception of four cases, the restricted OLS estimates using the aggregated data (Method I) present greater correlation coefficients than the unrestricted OLS estimates using the stacked district data (Method III). However the absolute values of these four unrestricted OLS probability estimates (Method III) present larger differences than the restricted OLS probability estimates obtained using the aggregate data. Moreover, although the t-values using stacked district data are significantly better than the corresponding t-values using aggregate data, a comparison between the time-patterns shows that the restricted OLS estimates obtained using the aggregate data still have patterns more similar to the observed point estimates

Fig. 8.1-8.13 Comparison between the Patterns of the OLS Estimates of the Transition Probabilities (Method III) and the Patterns of the Corresponding Observed Point Estimates

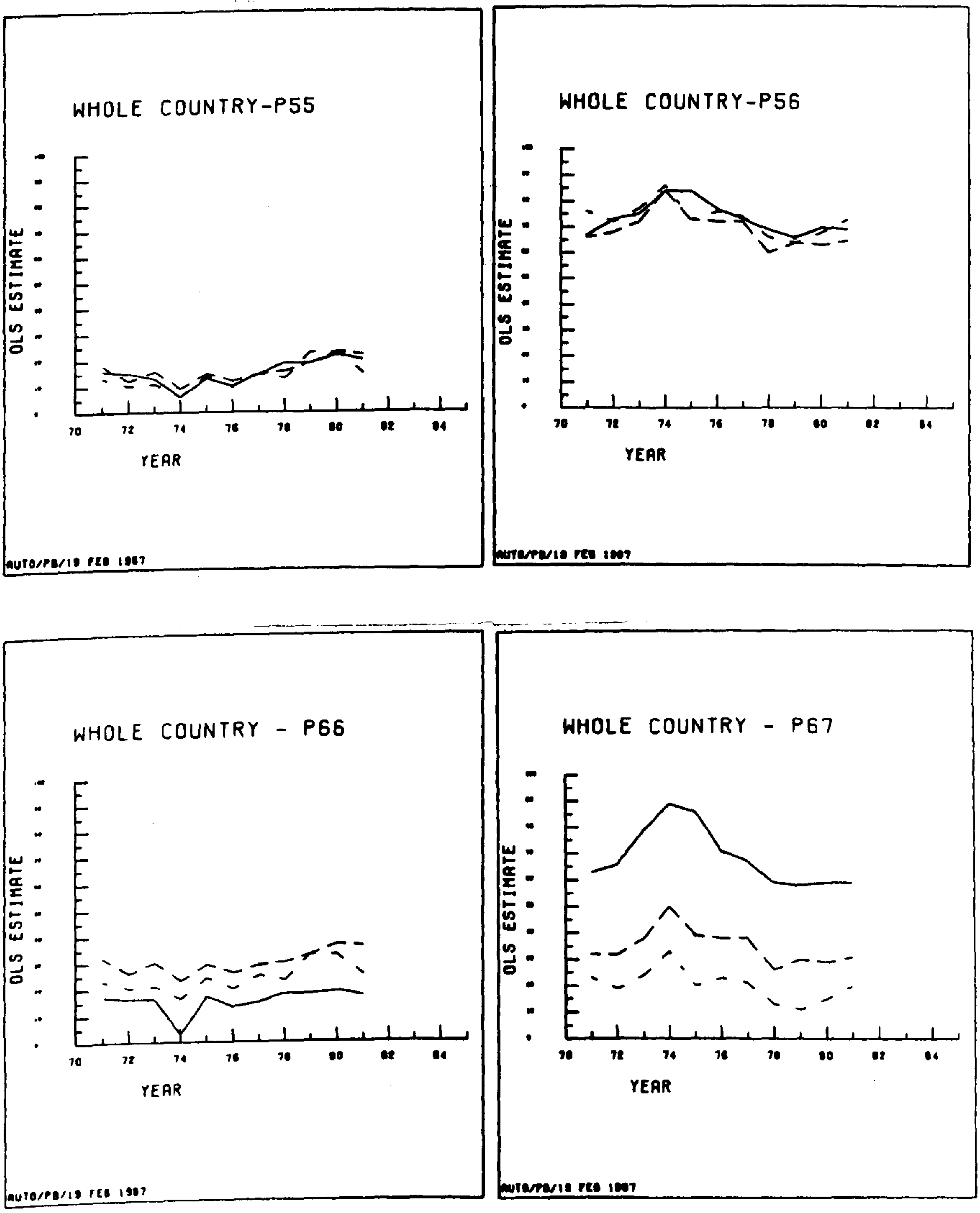


Fig. 8.1-8.13 Comparison between the Patterns of the OLS Estimates of the Transition Probabilities (Method III) and the Patterns of the Corresponding Observed Point Estimates (Continued)

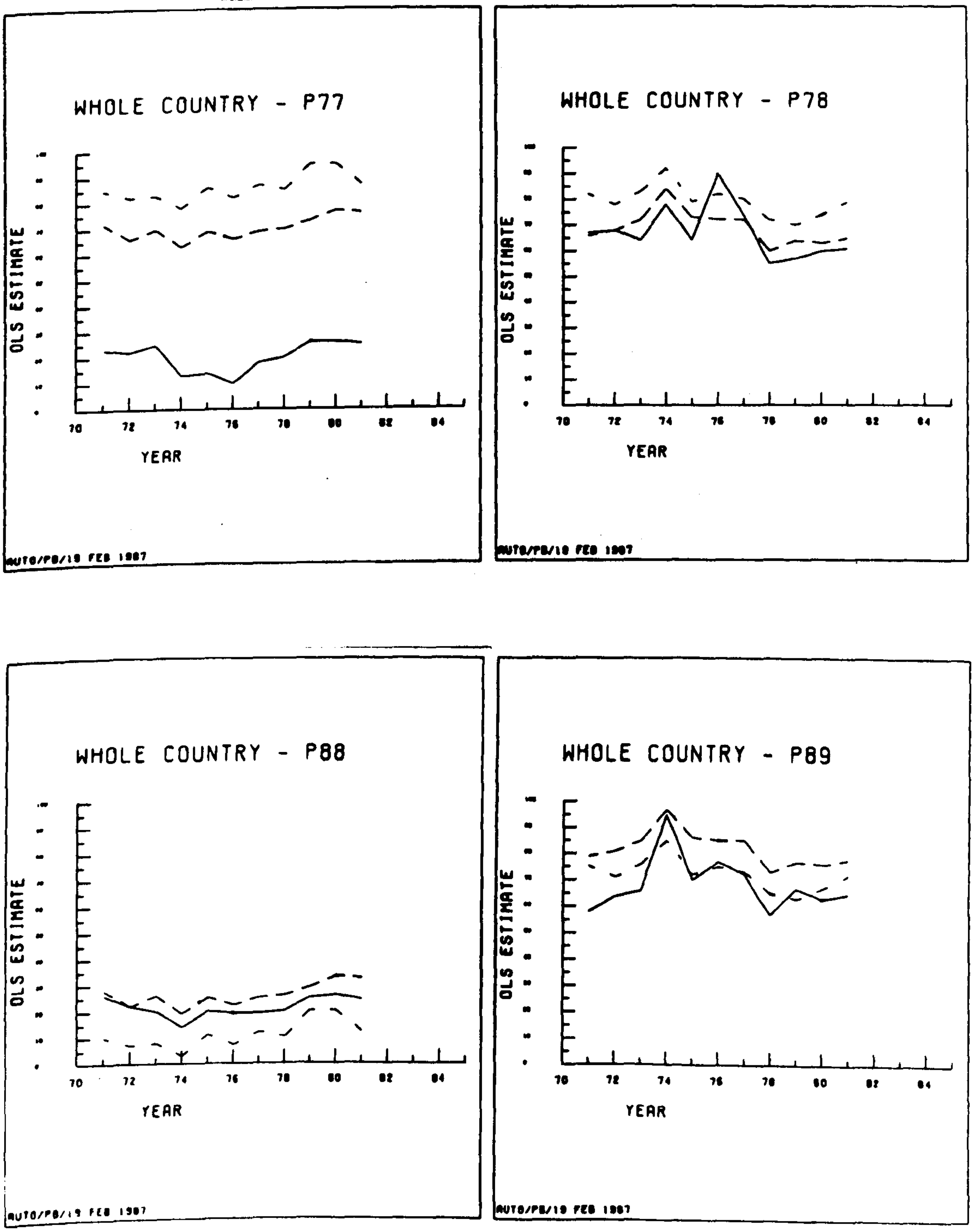
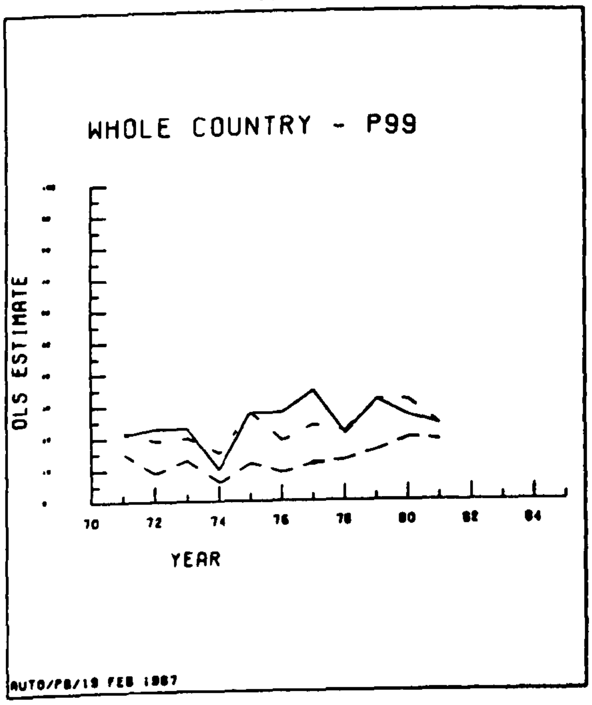
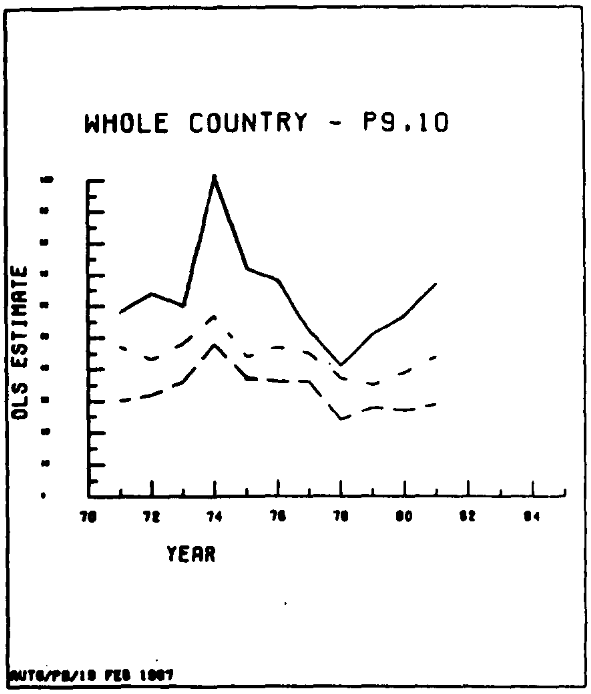
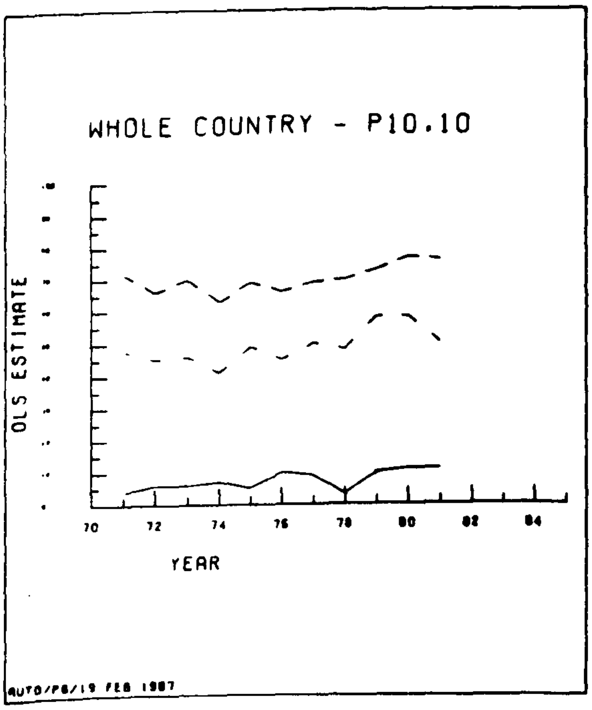


Fig. 8.1-8.13 Comparison between the Patterns of the OLS Estimates of the Transition Probabilities (Method III) and the Patterns of the Corresponding Observed Point Estimates (Continued)







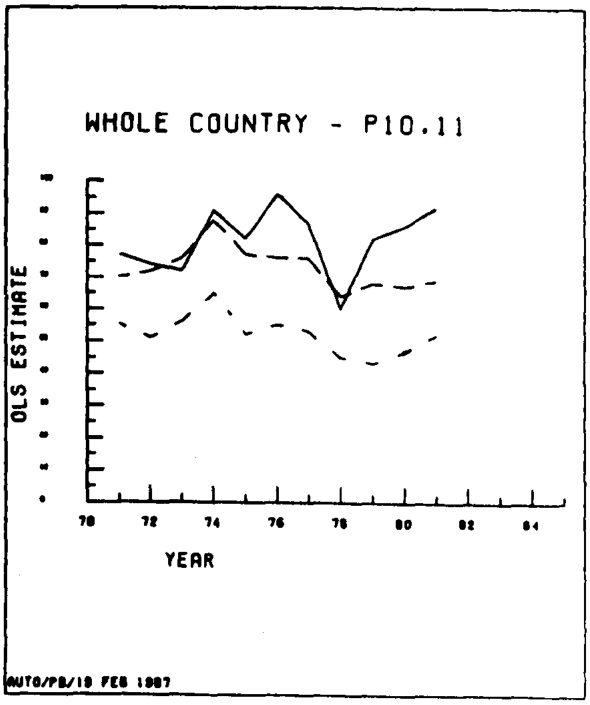
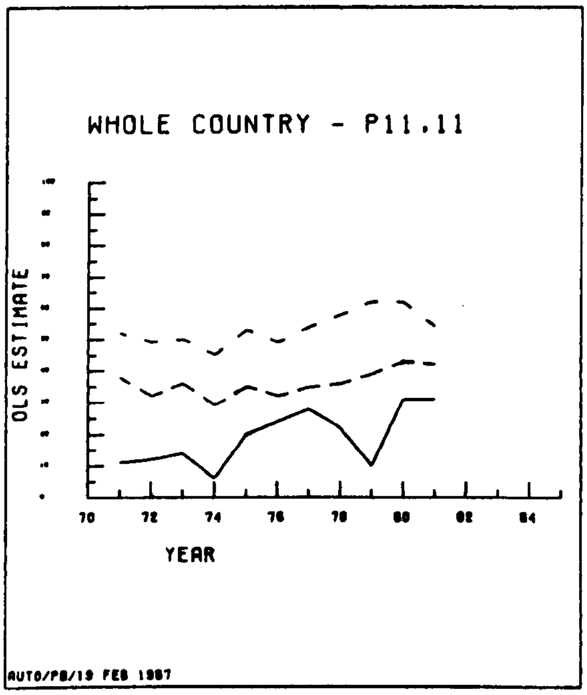


Fig. 8.1-8.13 **Comparison between the Patterns of the OLS Estimates of the Transition Probabilities (Method III) and the Patterns of the Corresponding Observed Point Estimates** (Continued)



——— Observed Point Estimates

— — — Unrestricted OLS Estimates

- - - - - Restricted OLS Estimates

TABLE 8.4: Comparison between the Estimates Obtained for the Stepwise Regression on Stacked District Data.

Transition Probability	Observed Point Estimate		Method III - Stepwise Regression on Stacked District Data (Whole Country)							
			Unrestricted OLS				Restricted OLS			
	Mean Value	Stand Dev.	Mean Value	Stand Dev.	R	D-W	Mean Value	Stand Dev.	R	D-W
P_{55}	.1527	.0469	.1603	.0433	.87977	1.80	.1398	.0508	.82191	1.80
P_{56}	.7336	.0622	.6901	.0667	.85360	2.29	.7300	.0593	.67097	1.36
P_{66}	.1536	.0441	.3003	.0433	.69935	2.20	.2398	.0508	.69330	2.51
P_{67}	.6873	.1123	.3501	.0667	.89875	2.71	.2000	.0593	.74608	1.51
P_{77}	.2018	.0569	.7003	.0433	.75253	.90*	.8598	.0508	.62597	.77*
P_{78}	.6709	.1023	.6901	.0667	.66957	1.95	.7900	.0593	.68511	1.70
P_{88}	.2136	.0367	.2603	.0433	.81894	1.08	.1098	.0508	.75105	.82
P_{89}	.6855	.1050	.8201	.0667	.89563	1.76	.7200	.0593	.73336	1.58
P_{99}	.2473	.0660	.1303	.0433	.38004	1.94	.2298	.0508	.63936	1.72
$P_{9,10}$	.6345	.1722	.3301	.0667	.85918	1.71	.4400	.0593	.82863	1.86
$P_{10,10}$	.0745	.0288	.7003	.0433	.42728	2.16	.4898	.0508	.46114	1.97
$P_{10,11}$	.8173	.1044	.7301	.0667	.43877	1.42	.5200	.0593	.41552	1.54
$P_{11,11}$	.1900	.0890	.3603	.0433	.53512	1.34	.5298	.0508	.39432	1.48

(a) R is the correlation coefficient between the observed point estimates and the transition probability estimates

$D\text{-}W_{0.05}$ (1.00; 1.32)

patterns than the unrestricted OLS estimates obtained using the stacked district data.

As previously described, the northern area of the country is more densely populated than the south and the coastal area more than the interior, which is the more rural area. These two distinct characteristics have suggested the division of the country into two regions. Region 1, the industrialized region, consists of the more developed areas which also include the more densely populated districts (Aveiro, Braga, Coimbra, Faro, Leiria, Lisboa, Porto, Santarem and Setubal), located in the coastal part of the country; Region 2, the rural region, consists of the less industrialized districts (Beja, Braganca, C. Branco, Evora, Guarda, Portalegre, V. Castelo, V. Real and Viseu), that is, the districts with appreciable rural characteristics and located in the inland part of the country. Figure 8.14 shows the geographical location of both regions.

The aim of this division is to improve the quality of the transition probability estimates, reasoning that the explanatory variables might differently affect the changes in the transition probabilities of each region. However, it is still assumed that within a region, the explanatory variables affect in the same way all the repetition probabilities and all the promotion probabilities, respectively.

Stepwise regressions using the SHAZAM program were now performed for both regions (Method IV) and the results are presented in Table 8.5 and Table 8.6. These tables show that only the unrestricted OLS estimation procedure for Region 2 gives significant t-values at the

Fig. 8.14 Geographical Location of the Regions Studied

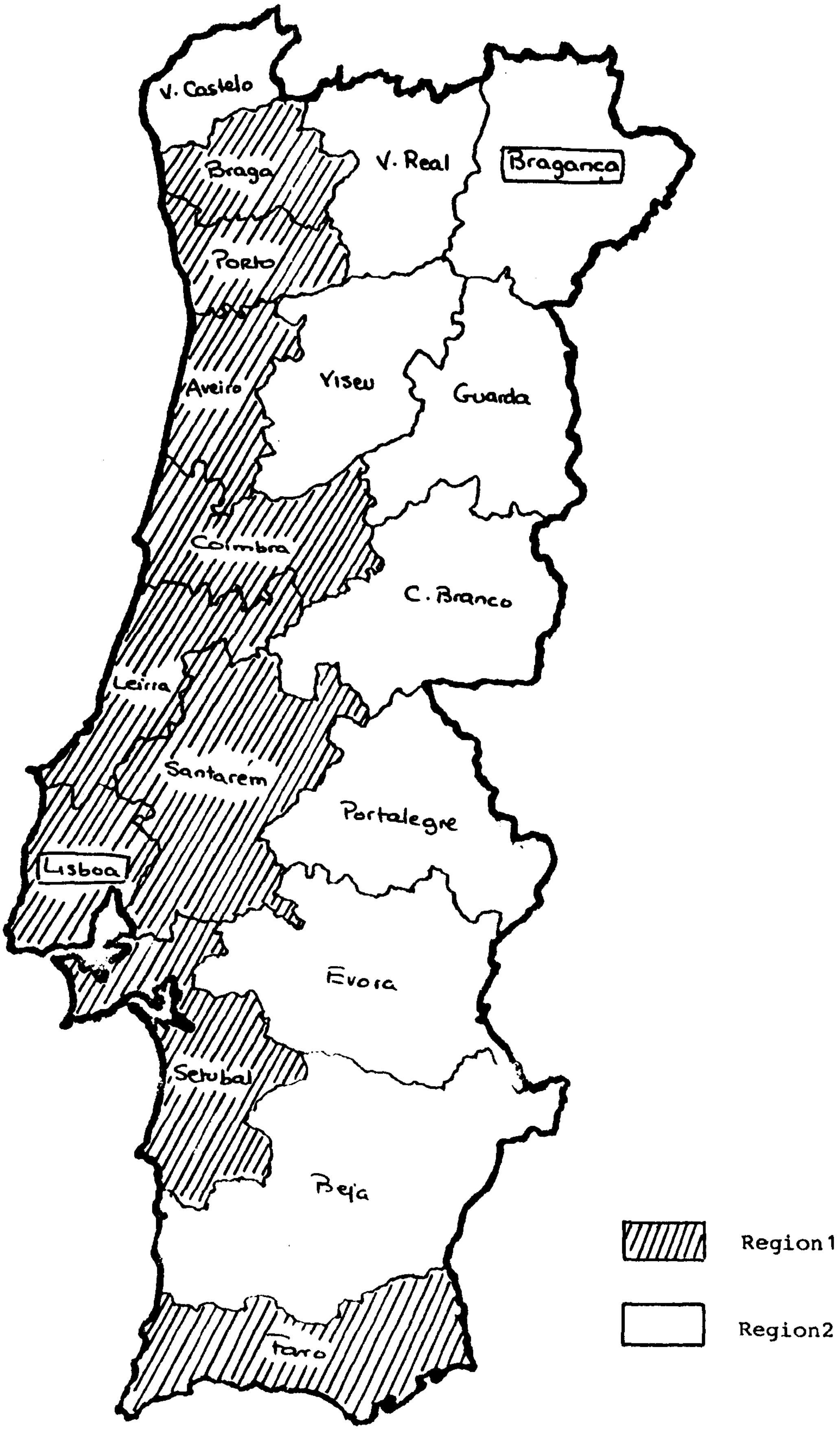


TABLE 8.5: Estimated Values of the a and δ Coefficients, Using Stacked District Data - Region 1 (Method IV).

Unrestricted OLS				Restricted OLS			
Coefficient	Estimated Value	St. Error	t-value	Coefficient	Estimated Value	St. Error	t-value
a_{55}	.16	.0095	15.55	a_{55}	.15	.0114	12.93
a_{56}	.77	.1015	7.63	a_{56}	.84	.1080	7.77
a_{66}	.21	.1242	1.66*	a_{66}	.11	.1323	.83*
a_{67}	.45	.0733	6.09	a_{67}	.26	.0787	3.35
a_{77}	.60	.0838	7.17	a_{77}	.79	.0900	8.78
a_{78}	.73	.0751	9.72	a_{78}	.81	.0915	8.91
a_{88}	.21	.0924	2.31	a_{88}	.09	.1127	.76*
a_{89}	.82	.0763	10.70	a_{89}	.66	.0901	7.30
a_{99}	.13	.0889	1.50*	a_{99}	.29	.1050	2.80
$a_{9,10}$	.36	.0868	4.17	$a_{9,10}$	.53	.0987	5.38
$a_{10,10}$	.67	.1390	4.74	$a_{10,10}$	.35	.1582	2.20
$a_{10,11}$	.75	.1309	5.70	$a_{10,11}$	.58	.1565	3.69
$a_{11,11}$	.34	.1371	2.45	$a_{11,11}$	.47	.1645	2.85
EDUC 1	.014	.0071	2.01	TEARN 1	-.0277	.0066	4.22
GDP 1	.038	.0086	4.43	BUS 1	-.0245	.0058	4.22
PCAP 2	-.036	.0099	3.64	HELP 1	.0310	.0056	5.58
TEARN 2	-.030	.0073	4.14	GDP 1	-.0339	.0100	3.40
HELP 2	.044	.0091	4.84	EARN 1	.0281	.0081	3.48
UNEMP 2	-.074	.0115	6.43	PCAP 1	-.0456	.0053	8.63
LFLEV 2	-.038	.0116	3.22	TEARN 2	.0349	.0072	4.86
				HELP 2	.0234	.0106	2.19
				UNQUAL 2	-.0254	.0099	2.56
				UNEMP 2	.0302	.0066	4.55

R^2 = .9591 $\qquad$ R^2 = .9930

$\bar{R}^2$ = .9579 $\qquad$ $\bar{R}^2$ = .9928

df = 672 $\qquad$ df = 768

TABLE 8.6: **Estimated Values of the a and δ Coefficients, Using Stacked District Data - Region 2 (Method IV)**

Unrestricted OLS				Restricted OLS			
Coefficient	Estimated Value	St. Error	t-value	Coefficient	Estimated Value	St. Error	t-value
a_{55}	.16	.0060	27.11	a_{55}	.13	.0072	17.77
a_{56}	.32	.0833	3.83	a_{56}	.41	.0553	7.36
a_{66}	.69	.0977	7.01	a_{66}	.54	.0707	7.68
a_{67}	.25	.0474	5.28	a_{67}	.21	.0377	5.50
a_{77}	.78	.0516	15.05	a_{77}	.75	.0571	13.40
a_{78}	.54	.0512	10.58	a_{78}	.66	.0576	11.40
a_{88}	.39	.0622	6.31	a_{88}	.25	.0697	3.59
a_{89}	.73	.0505	14.53	a_{89}	.96	.0625	15.34
a_{99}	.21	.0587	3.53	a_{99}	.03	.0655	.52*
$a_{9,10}$	.24	.0654	3.66	$a_{9,10}$	.13	.0545	2.41
$a_{10,10}$	.73	.1010	7.20	$a_{10,10}$	.84	.1037	8.14
$a_{10,11}$	.80	.0778	10.33	$a_{10,11}$	.88	.1112	7.91
$a_{11,11}$	.29	.0804	3.58	$a_{11,11}$	.21	.1186	1.78*
TEARN 1	.021	.0066	3.17	EARN 1	-.0115	.0045	2.58
BUS 1	-.024	.0090	2.63	POPLEV 1	.0405	.0070	5.80
HELP 1	-.018	.0077	2.30	TEARN 2	.0166	.0030	5.47
GDP 1	.045	.0091	4.89	UNEMP 2	-.0187	.0049	3.77
EARN 1	-.017	.0057	2.90	LFLEV 2	.1315	.0154	8.55
PCAP 2	-.053	.0068	7.84	POPLEV 2	-.1588	.0176	9.05
TEARN 2	-.048	.0066	7.36				
HELP 2	.080	.0083	9.60				
UNQUAL 2	.050	.0102	4.92				
UNEMP 2	-.084	.0101	8.31				

R^2 = .9884 $\qquad$ R^2 = .9953

$\bar{R}^2$ = .9881 $\qquad$ $\bar{R}^2$ = .9952

df = 669 $\qquad$ df = 772

5% level of significance for the set of the parameters of the model. The remaining estimation procedures present non-significant t-values for two 'a' coefficients.

It must be noted from the tables that the unrestricted OLS repetition probability estimates for Region 1 are equal for all districts as only EDUC and GDP have been the explanatory variables selected to determine the changes in the repetition probabilities; as already noted in section 8.2, these explanatory variables have, by their nature, the same national value for all districts.

The transition probability estimates and the corresponding t-values are presented in Tables 8.7 to 8.10 for one district of each region. The selected districts are Lisboa for Region 1 and Braganca for Region 2, the two sample districts usually used by the researchers and the Portuguese governmental education departments when micro studies are performed. It can be seen that the attempt at dividing the country into regions has not generated better transition probability estimates than the estimates obtained when stacking all district data. The OLS estimates $\hat{p}_{67}$, $\hat{p}_{77}$, $\hat{p}_{9,10}$ and $\hat{p}_{10,10}$ are completely distorted from the reality for both regions, as very high repetition probability estimates and very low promotion probability estimates have occured. Furthermore, the promotion probability estimate $\hat{p}_{56}$ has also very low values for Region 2.

The transition probability estimates for Lisboa and Braganca obtained by applying stepwise regression to the stacked district data (Method III) are presented in Tables 8.11 to 8.14, and a comparison between some statistics obtained after performing the regression of the observed point estimates on the stepwise regression estimates,

TABLE 8.7: Method IV - Transition Probability Estimates Using Stepwise Regression Applied to Region 1 with all the Explanatory Variables (Unrestricted OLS) - LISBOA

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.11	..79	.16	.47	.55	.75	.16	.84	.08	.38	.62	.77	.29
	(7.19)	(7.89)	(1.27)	(6.01)	(6.51)	(9.72)	(1.71)	(10.04)	(.90)	(4.13)	(4.30)	(5.67)	(2.09)
1972	.13	.81	.18	.49	.57	.77	.18	.86	.10	.40	.64	.79	.31
	(10.13)	(7.97)	(1.49)	(6.38)	(6.83)	(9.65)	(2.03)	(11.28)	(1.15)	(4.16)	(4.56)	(6.02)	(2.28)
1973	.15	.83	.20	.51	.59	.79	.20	.88	.12	.42	.66	.81	.33
	(12.86)	(8.44)	(1.66)	(6.69)	(7.06)	(10.80)	(2.26)	(10.51)	(1.36)	(4.53)	(4.76)	(5.82)	(2.43)
1974	.12	.89	.17	.57	.56	.85	.17	.94	.09	.48	.63	.87	.30
	(9.36)	(8.67)	(1.38)	(7.79)	(6.71)	(10.56)	(1.85)	(13.58)	(1.08)	(4.84)	(4.48)	(6.58)	(2.20)
1975	.16	.76	.21	.44	.60	.72	.21	.81	.13	.35	.67	.74	.34
	(7.58)	(7.23)	(1.74)	(5.06)	(6.75)	(9.80)	(2.31)	(9.36)	(1.28)	(3.13)	(4.65)	(5.98)	(2.38)
1976	.12	.82	.17	.50	.56	.78	.17	.87	.09	.41	.63	.80	.30
	(9.34)	(7.96)	(1.38)	(6.45)	(6.70)	(10.69)	(1.85)	(10.49)	(1.06)	(4.24)	(4.45)	(5.99)	(2.20)
1977	.15	.81	.20	.49	.59	.77	.20	.86	.12	.40	.66	.79	.33
	(14.54)	(7.62)	(1.62)	(6.78)	(7.06)	(8.64)	(2.19)	(11.59)	(1.35)	(4.31)	(4.74)	(5.72)	(2.41)
1978	.16	.67	.21	.35	.60	.63	.21	.72	.13	.26	.67	.65	.34
	(14.77)	(6.46)	(1.65)	(4.72)	(7.08)	(7.64)	(2.17)	(8.18)	(1.47)	(3.39)	(4.76)	(4.64)	(2.45)
1979	.19	.69	.24	.37	.63	.65	.24	.74	.16	.28	.70	.67	.37
	(16.90)	(6.58)	(1.95)	(4.98)	(7.37)	(7.91)	(2.51)	(8.58)	(1.82)	(3.45)	(5.06)	(4.85)	(2.69)
1980	.23	.70	.28	.38	.67	.66	.28	.75	.20	.29	.74	.68	.48
	(14.36)	(6.84)	(2.26)	(5.03)	(7.52)	(8.24)	(2.84)	(8.60)	(2.07)	(3.26)	(5.31)	(4.77)	(2.89)
1981	.21	.70	.26	.38	.65	.66	.26	.75	.18	.29	.72	.68	.39
	(13.72)	(6.86)	(2.07)	(5.01)	(7.36)	(7.80)	(2.58)	(8.75)	(1.98)	(3.00)	(5.12)	(4.74)	(2.77)

Figures in parenthesis are t-values

TABLE 8.8: Method IV - Transition Probability Estimates Using Stepwise Regression Applied to Region I with all Explanatory Variables (Restricted OLS) - LISBOA

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.14	.85	.10	.27	.78	.82	.08	.67	.28	.54	.33	.59	.46
	(6.68)	(7.96)	(.72)	(3.21)	(8.48)	(8.77)	(.68)	(7.27)	(2.47)	(5.61)	(1.99)	(3.72)	(2.79)
1972	.12	.86	.08	.28	.76	.83	.06	.68	.26	.55	.31	.60	.44
	(7.81)	(7.98)	(.58)	(3.42)	(8.58)	(9.10)	(.51)	(7.90)	(2.44)	(5.37)	(1.89)	(3.94)	(2.72)
1973	.10	.85	.06	.27	.74	.82	.04	.67	.24	.54	.29	.59	.42
	(6.80)	(8.10)	(.43)	(3.34)	(8.44)	(10.03)	(.33)	(7.16)	(2.31)	(5.52)	(1.76)	(3.66)	(2.58)
1974	.14	1.04	.10	.48	.78	1.01	.08	.86	.28	.73	.33	.78	.46
	(7.05)	(9.49)	(.79)	(5.81)	(7.91)	(11.17)	(.70)	(10.09)	(2.48)	(6.74)	(2.08)	(5.27)	(2.37)
1975	.17	.78	.13	.20	.81	.75	.11	.60	.31	.47	.33	.52	.49
	(12.22)	(7.34)	(.99)	(2.23)	(8.53)	(8.68)	(.96)	(6.47)	(2.83)	(4.11)	(2.27)	(3.42)	(2.72)
1976	.18	.87	.14	.29	.82	.84	.12	.69	.32	.56	.37	.61	.50
	(15.20)	(8.09)	(1.10)	(3.67)	(8.82)	(9.40)	(1.10)	(7.65)	(2.99)	(5.13)	(2.40)	(3.86)	(2.93)
1977	.16	.86	.12	.28	.80	.83	.10	.68	.30	.55	.35	.60	.48
	(13.72)	(7.63)	(.89)	(3.70)	(8.93)	(7.29)	(.88)	(8.15)	(2.85)	(5.22)	(2.22)	(3.58)	(3.01)
1978	.15	.76	.11	.18	.79	.73	.09	.58	.29	.45	.34	.50	.42
	(12.84)	(6.92)	(.80)	(2.31)	(8.84)	(7.40)	(.76)	(5.61)	(2.76)	(5.17)	(2.12)	(3.04)	(2.93)
1979	.16	.79	.12	.21	.80	.76	.10	.61	.30	.48	.35	.53	.48
	(11.61)	(7.08)	(.87)	(2.71)	(8.84)	(8.00)	(.84)	(5.89)	(2.85)	(4.63)	(2.18)	(3.34)	(3.07)
1980	.15	.77	.11	.19	.79	.74	.09	.59	.29	.46	.34	.51	.47
	(11.87)	(7.05)	(.85)	(2.36)	(8.86)	(7.83)	(.82)	(5.85)	(2.84)	(4.53)	(2.17)	(3.04)	(3.06)
1981	.19	.80	.15	.22	.83	.77	.13	.62	.33	.49	.38	.54	.51
	(14.26)	(7.40)	(9.90)	(2.84)	(8.92)	(8.10)	(1.10)	(6.61)	(3.05)	(4.96)	(2.41)	(3.33)	(3.12)

Figures in parenthesis are t-values

TABLE 8.9: Method IV - Transition Probability Estimates Using Stepwise Regression Applied to Region 2 with all the Explanatory Variables (Unrestricted OLS) - BRAGANCA

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.19	.32	.72	.25	.81	.54	.42	.73	.24	.24	.76	.80	.32
	(11.84)	(3.92)	(7.25)	(4.68)	(14.02)	(10.01)	(6.24)	(13.21)	(3.56)	(3.45)	(7.29)	(9.52)	(3.65)
1972	.19	.33	.72	.26	.81	.55	.42	.74	.24	.25	.76	.81	.32
	(12.36)	(3.93)	(7.37)	(5.03)	(14.07)	(9.76)	(6.26)	(14.53)	(3.72)	(3.44)	(7.39)	(10.01)	(3.75)
1973	.19	.30	.71	.23	.81	.52	.42	.71	.24	.22	.76	.78	.32
	(18.12)	(3.78)	(7.34)	(4.45)	(14.58)	(10.61)	(7.00)	(12.68)	(3.78)	(3.34)	(7.55)	(9.03)	(3.90)
1974	.13	.46	.66	.39	.75	.68	.36	.87	.18	.38	.70	.94	.26
	(10.85)	(5.50)	(6.73)	(8.12)	(13.51)	(12.24)	(5.60)	(18.46)	(2.93)	(5.03)	(6.48)	(11.39)	(3.08)
1975	.16	.27	.69	.20	.78	.49	.39	.68	.21	.19	.73	.75	.29
	(8.54)	(3.07)	(6.85)	(3.53)	(14.49)	(8.20)	(5.14)	(12.19)	(3.14)	(2.09)	(6.62)	(10.60)	(3.21)
1976	.11	.37	.64	.30	.73	.59	.34	.78	.16	.29	.68	.85	.24
	(7.89)	(4.37)	(6.27)	(6.12)	(13.72)	(11.38)	(5.46)	(15.05)	(2.45)	(4.31)	(6.28)	(11.11)	(2.81)
1977	.11	.36	.64	.29	.73	.58	.34	.77	.16	.28	.68	.84	.24
	(7.17)	(4.16)	(6.26)	(5.96)	(12.54)	(9.66)	(5.56)	(12.98)	(2.31)	(4.13)	(6.11)	(9.18)	(2.73)
1978	.15	.20	.68	.13	.77	.42	.38	.61	.20	.12	.72	.68	.28
	(18.49)	(2.37)	(7.00)	(2.62)	(14.57)	(7.82)	(6.04)	(9.87)	(3.38)	(1.94)	(6.99)	(7.83)	(3.46)
1979	.18	.26	.71	.19	.80	.48	.41	.67	.23	.18	.75	.74	.31
	(20.17)	(3.09)	(7.26)	(4.09)	(14.76)	(8.81)	(6.06)	(11.80)	(3.71)	(2.88)	(7.42)	(9.11)	(3.71)
1980	.18	.26	.71	.19	.80	.48	.41	.67	.23	.18	.75	.74	.31
	(16.08)	(3.08)	(7.21)	(3.90)	(14.39)	(8.98)	(5.79)	(11.31)	(3.62)	(2.79)	(7.28)	(8.71)	(3.62)
1981	.17	.40	.70	.33	.79	.62	.40	.81	.22	.32	.74	.88	.30
	(15.49)	(4.63)	(7.11)	(6.67)	(14.06)	(10.76)	(5.68)	(14.33)	(3.44)	(4.32)	(7.27)	(10.40)	(3.51)

Figures in parenthesis are t-values

TABLE 8.10: Method IV - Transition Probability Estimates using Stepwise Regression Applied to Region 2 with all the Explanatory Variables (Restricted OLS) - BRAGANCA

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.09 (7.61)	.42 (7.60)	.51 (6.96)	.22 (5.33)	.72 (11.89)	.67 (10.96)	.22 (3.17)	.97 (15.98)	.00 (.06)	.14 (2.15)	.81 (7.19)	.89 (7.85)	.18 (1.47)
1972	.10 (8.22)	.44 (7.96)	.52 (7.02)	.24 (6.07)	.73 (12.00)	.69 (11.69)	.23 (3.24)	.99 (16.33)	.01 (.12)	.16 (2.59)	.82 (7.25)	.91 (8.11)	.19 (1.51)
1973	.10 (9.14)	.43 (7.87)	.52 (7.11)	.23 (6.03)	.73 (12.14)	.68 (11.65)	.23 (3.34)	.98 (16.47)	.01 (.21)	.15 (2.52)	.82 (2.33)	.90 (8.00)	.19 (1.56)
1974	.08 (6.78)	.47 (8.45)	.50 (7.07)	.27 (7.31)	.71 (12.19)	.72 (12.43)	.21 (2.87)	1.02 (16.50)	-.01 (.21)	.19 (3.43)	.80 (7.63)	.94 (8.28)	.17 (1.37)
1975	.07 (5.56)	.51 (9.12)	.49 (7.23)	.31 (7.60)	.70 (12.29)	.76 (14.14)	.20 (2.71)	1.06 (14.96)	-.02 (.25)	.23 (3.90)	.79 (8.08)	.98 (9.68)	.16 (1.36)
1976	.10 (12.12)	.49 (7.99)	.52 (7.49)	.24 (6.39)	.73 (12.92)	.69 (12.32)	.23 (3.28)	.99 (15.66)	.01 (.16)	.16 (2.92)	.82 (8.08)	.91 (8.39)	.19 (1.62)
1977	.12 (16.73)	.38 (6.71)	.54 (7.64)	.18 (4.79)	.75 (13.15)	.63 (10.02)	.25 (3.59)	.93 (14.45)	.03 (.46)	.10 (2.19)	.84 (8.12)	.85 (7.23)	.21 (1.77)
1978	.12 (16.89)	.44 (7.47)	.54 (7.63)	.24 (5.91)	.75 (13.13)	.69 (11.03)	.25 (3.62)	.99 (13.63)	.03 (.49)	.16 (3.51)	.84 (8.08)	.91 (7.95)	.21 (1.78)
1979	.16 (18.04)	.32 (5.68)	.58 (8.08)	.12 (2.87)	.79 (13.58)	.57 (9.00)	.29 (4.06)	.87 (13.14)	.07 (1.06)	.04 (.67)	.88 (8.54)	.79 (6.78)	.25 (2.11)
1980	.17 (16.64)	.37 (6.44)	.59 (8.19)	.17 (4.18)	.80 (13.63)	.62 (10.22)	.30 (4.14)	.92 (13.21)	.08 (1.20)	.09 (1.60)	.89 (8.67)	.84 (7.23)	.26 (2.20)
1981	.19 (14.24)	.36 (6.33)	.61 (8.36)	.16 (3.75)	.82 (13.56)	.61 (10.33)	.32 (4.26)	.91 (13.44)	.10 (1.45)	.08 (1.12)	.91 (8.83)	.83 (6.98)	.28 (2.35)

Figures in parenthesis are t-values

TABLE 8.11: Method III - Transition Probability Estimates using Stepwise Regression on Stacked District Data (All Districts) - Unrestricted OLS -- LISBOA

YEAR	p_{55}	p_{56}	p_{66}	p_{67}	p_{77}	p_{78}	p_{88}	p_{89}	p_{99}	$p_{9,10}$	$p_{10,10}$	$p_{10,11}$	$p_{11,11}$
1971	.16 (9.75)	.70 (10.19)	.30 (3.37)	.36 (6.96)	.70 (11.41)	.70 (12.63)	.26 (3.69)	.83 (13.94)	.13 (1.91)	.34 (5.51)	.70 (6.93)	.74 (7.70)	.36 (3.58)
1972	.15 (14.60)	.72 (10.33)	.29 (3.42)	.38 (7.64)	.69 (12.30)	.72 (12.71)	.25 (3.88)	.85 (15.85)	.12 (1.91)	.36 (5.39)	.69 (7.14)	.76 (8.24)	.35 (3.65)
1973	.15 (18.25)	.76 (11.28)	.29 (3.47)	.42 (8.42)	.69 (12.62)	.76 (14.81)	.25 (3.93)	.89 (15.02)	.12 (1.92)	.40 (6.16)	.69 (7.17)	.80 (8.24)	.35 (3.67)
1974	.07 (4.49)	.88 (12.39)	.21 (2.46)	.54 (10.33)	.61 (10.39)	.88 (15.30)	.17 (2.39)	1.01 (16.81)	.04 (.66)	.52 (7.08)	.61 (5.88)	.92 (9.57)	.27 (2.67)
1975	.15 (9.66)	.68 (9.61)	.29 (3.43)	.34 (6.17)	.69 (11.80)	.68 (13.20)	.25 (3.75)	.81 (13.57)	.12 (1.61)	.32 (4.31)	.69 (6.78)	.72 (8.41)	.35 (3.44)
1976	.13 (13.22)	.74 (10.43)	.27 (3.08)	.40 (7.78)	.67 (11.98)	.74 (13.44)	.23 (3.37)	.87 (11.98)	.10 (1.56)	.38 (5.68)	.67 (6.84)	.78 (8.17)	.33 (3.38)
1977	.16 (20.17)	.69 (9.50)	.30 (3.55)	.35 (7.36)	.70 (12.70)	.69 (10.50)	.26 (4.05)	.82 (15.85)	.13 (2.13)	.33 (5.05)	.70 (7.33)	.73 (7.40)	.36 (3.80)
1978	.16 (21.55)	.57 (7.96)	.30 (3.49)	.23 (4.58)	.70 (12.67)	.57 (9.43)	.26 (3.88)	.70 (11.22)	.13 (2.16)	.21 (3.97)	.70 (7.27)	.61 (6.10)	.36 (3.76)
1979	.19 (23.24)	.62 (8.66)	.33 (3.84)	.28 (5.82)	.73 (13.05)	.62 (10.58)	.29 (4.24)	.75 (12.00)	.16 (2.52)	.26 (4.02)	.73 (7.50)	.66 (6.81)	.39 (3.99)
1980	.23 (19.24)	.61 (8.79)	.37 (4.32)	.27 (5.58)	.77 (13.25)	.61 (10.82)	.33 (4.76)	.74 (12.08)	.20 (2.00)	.25 (3.83)	.77 (7.74)	.65 (6.59)	.43 (4.29)
1981	.21 (18.96)	.62 (8.85)	.35 (4.02)	.28 (5.64)	.75 (12.97)	.62 (10.38)	.31 (4.36)	.75 (12.55)	.18 (2.75)	.26 (3.91)	.75 (7.56)	.66 (6.60)	.41 (4.11)

Figures in parenthesis are t-values

TABLE 8.12: Method III - Transition Probability Estimates Using Stepwise Regression on Stacked District Data (All Districts) - Restricted OLS - LISBOA

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.12	.78	.22	.25	.84	.84	.09	.77	.21	.49	.47	.57	.51
	(5.79)	(10.74)	(2.34)	(4.38)	(13.17)	(11.69)	(1.01)	(11.85)	(2.52)	(6.32)	(4.02)	(5.06)	(4.26)
1972	.12	.74	.22	.21	.84	.80	.09	.73	.21	.45	.47	.53	.51
	(6.27)	(10.08)	(2.46)	(3.68)	(13.22)	(11.40)	(1.08)	(10.86)	(2.59)	(6.16)	(4.26)	(4.58)	(4.19)
1973	.10	.77	.20	.24	.82	.83	.07	.76	.19	.48	.45	.56	.49
	(5.60)	(10.63)	(2.22)	(4.44)	(13.09)	(12.74)	(.85)	(11.07)	(2.39)	(6.09)	(3.71)	(4.72)	(3.78)
1974	.04	.94	.14	.41	.76	1.00	.01	.93	.13	.65	.39	.73	.43
	(2.61)	(13.05)	(1.58)	(8.14)	(13.03)	(14.99)	(.10)	(14.82)	(1.69)	(8.74)	(3.48)	(7.04)	(3.60)
1975	.12	.73	.22	.20	.84	.79	.09	.72	.21	.44	.47	.52	.51
	(10.51)	(10.38)	(2.62)	(3.69)	(14.43)	(14.21)	(1.27)	(10.42)	(2.63)	(5.96)	(4.59)	(5.77)	(4.44)
1976	.09	.80	.19	.27	.81	.86	.06	.79	.18	.51	.44	.59	.48
	(5.14)	(11.05)	(2.11)	(5.04)	(13.35)	(12.67)	(.70)	(10.96)	(2.15)	(5.96)	(3.68)	(5.26)	(4.34)
1977	.20	.71	.30	.18	.92	.77	.17	.70	.29	.42	.55	.50	.59
	(17.66)	(9.85)	(3.45)	(3.95)	(15.75)	(11.09)	(2.48)	(11.79)	(3.89)	(7.20)	(5.52)	(4.34)	(5.16)
1978	.14	.62	.24	.09	.86	.68	.11	.61	.23	.33	.49	.41	.53
	(10.24)	(8.51)	(2.72)	(1.85)	(14.63)	(9.73)	(1.42)	(9.21)	(3.65)	(5.52)	(4.65)	(3.60)	(4.82)
1979	.22	.61	.32	.08	.94	.67	.19	.60	.31	.32	.55	.40	.61
	(16.09)	(8.37)	(3.58)	(1.64)	(15.17)	(9.43)	(2.14)	(8.82)	(4.52)	(4.84)	(5.31)	(3.42)	(5.14)
1980	.22	.64	.32	.11	.94	.70	.19	.63	.31	.35	.55	.43	.61
	(20.29)	(8.85)	(3.59)	(2.22)	(15.72)	(10.74)	(2.35)	(9.32)	(4.14)	(5.17)	(5.67)	(3.68)	(5.14)
1981	.16	.69	.26	.16	.88	.75	.13	.68	.25	.40	.17	.48	.55
	(15.23)	(9.56)	(2.93)	(3.26)	(15.60)	(11.91)	(1.76)	(10.43)	(3.45)	(6.29)	(5.13)	(4.28)	(5.18)

Figures in parenthesis are t-values

TABLE 8.13: Method III - Transition Probability Estimates Using Stepwise Regression on Stacked District Data (All Districts) - Unrestricted OLS - BRAGANCA

YEAR	p_{55}	p_{56}	p_{66}	p_{67}	p_{77}	p_{78}	p_{88}	p_{89}	p_{99}	$p_{9,10}$	$p_{10,10}$	$p_{10,11}$	$p_{11,11}$
1971	.13 (9.64)	.64 (9.33)	.27 (3.08)	.30 (5.78)	.67 (11.42)	.64 (11.01)	.23 (3.37)	.77 (12.64)	.10 (1.55)	.28 (4.43)	.67 (6.77)	.68 (6.93)	.33 (3.36)
1972	.16 (17.23)	.70 (9.96)	.30 (3.55)	.36 (7.09)	.70 (12.68)	.70 (11.47)	.26 (4.05)	.83 (14.87)	.13 (2.08)	.34 (4.83)	.70 (7.26)	.74 (7.67)	.36 (3.77)
1973	.17 (12.34)	.70 (10.25)	.31 (3.55)	.36 (7.14)	.71 (11.99)	.70 (12.65)	.27 (4.00)	.83 (13.80)	.14 (2.01)	.34 (5.16)	.71 (7.17)	.74 (7.39)	.37 (3.73)
1974	.06 (4.96)	.82 (11.65)	.20 (2.27)	.48 (9.76)	.60 (10.37)	.82 (14.51)	.16 (2.33)	.95 (17.51)	.03 (.48)	.46 (6.70)	.60 (6.11)	.86 (9.21)	.26 (2.66)
1975	.14 (9.08)	.73 (10.18)	.28 (3.35)	.39 (6.74)	.68 (11.47)	.73 (13.59)	.24 (3.70)	.86 (14.37)	.11 (1.56)	.37 (4.40)	.68 (6.78)	.77 (8.85)	.34 (3.41)
1976	.14 (14.38)	.73 (10.37)	.28 (3.25)	.39 (7.63)	.68 (12.25)	.73 (13.51)	.24 (3.57)	.86 (15.25)	.11 (1.78)	.37 (5.82)	.68 (6.96)	.77 (8.32)	.34 (3.51)
1977	.15 (17.57)	.70 (9.60)	.29 (3.45)	.36 (7.54)	.69 (12.64)	.70 (11.04)	.25 (3.82)	.83 (15.57)	.12 (1.94)	.34 (5.28)	.69 (7.11)	.74 (7.50)	.35 (3.64)
1978	.16 (17.11)	.59 (8.37)	.30 (3.50)	.25 (5.05)	.70 (12.62)	.59 (10.31)	.26 (3.80)	.72 (11.75)	.13 (2.12)	.23 (4.12)	.70 (7.14)	.63 (6.48)	.36 (3.70)
1979	.19 (19.41)	.67 (9.41)	.33 (3.93)	.33 (6.83)	.73 (13.00)	.67 (11.93)	.29 (4.27)	.80 (13.30)	.16 (2.60)	.31 (4.65)	.73 (7.44)	.71 (7.59)	.39 (4.01)
1980	.23 (18.87)	.63 (9.14)	.37 (4.44)	.29 (6.03)	.77 (13.42)	.63 (11.70)	.33 (4.83)	.76 (12.28)	.20 (3.01)	.27 (4.34)	.77 (7.74)	.67 (6.90)	.43 (4.33)
1981	.21 (17.27)	.69 (9.64)	.35 (4.10)	.35 (6.91)	.75 (13.01)	.69 (11.30)	.31 (4.37)	.82 (12.53)	.18 (2.81)	.33 (5.57)	.75 (7.50)	.73 (6.99)	.41 (4.12)

Figures in parenthesis are t-values

TABLE 8.14: Method III - Transition Probability Estimates Using Stepwise Regression on Stacked District Data (All Districts) - Restricted OLS - BRAGANCA

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.07	.76	.17	.23	.79	.82	.04	.75	.16	.47	.42	.55	.49
	(3.50)	(10.58)	(1.80)	(4.15)	(12.78)	(11.05)	(.42)	(11.95)	(2.03)	(5.91)	(3.61)	(4.83)	(3.83)
1972	.14	.72	.24	.19	.86	.78	.11	.71	.23	.43	.49	.51	.53
	(7.63)	(9.95)	(2.68)	(3.46)	(13.55)	(11.05)	(1.29)	(11.14)	(2.87)	(5.89)	(4.45)	(4.40)	(4.40)
1973	.13	.74	.23	.21	.85	.80	.10	.73	.22	.45	.48	.53	.52
	(7.81)	(10.33)	(2.60)	(4.02)	(13.19)	(11.80)	(1.28)	(11.81)	(2.96)	(6.13)	(4.12)	(4.56)	(4.06)
1974	.05	.83	.15	.30	.77	.89	.02	.82	.14	.54	.40	.62	.44
	(3.09)	(11.60)	(1.74)	(6.12)	(12.75)	(13.67)	(.24)	(13.38)	(1.78)	(7.74)	(3.65)	(5.61)	(3.58)
1975	.16	.73	.26	.20	.88	.79	.13	.72	.25	.44	.51	.52	.55
	(10.89)	(10.22)	(3.10)	(3.54)	(14.37)	(14.72)	(1.90)	(9.68)	(2.91)	(6.03)	(5.15)	(5.70)	(4.53)
1976	.15	.76	.25	.23	.87	.82	.12	.75	.24	.47	.50	.55	.54
	(11.72)	(10.68)	(2.79)	(4.56)	(15.37)	(13.82)	(1.60)	(11.32)	(2.86)	(6.56)	(4.60)	(5.20)	(4.91)
1977	.14	.72	.24	.19	.86	.78	.11	.71	.23	.43	.49	.51	.53
	(11.82)	(9.83)	(2.79)	(4.05)	(14.69)	(10.96)	(1.51)	(11.47)	(3.24)	(7.22)	(4.58)	(4.36)	(4.81)
1978	.12	.65	.22	.12	.84	.71	.09	.64	.21	.36	.47	.44	.51
	(8.03)	(8.82)	(2.45)	(2.31)	(14.23)	(9.95)	(1.10)	(9.55)	(3.18)	(5.70)	(4.18)	(3.83)	(4.79)
1979	.21	.66	.31	.13	.93	.72	.18	.65	.30	.37	.56	.45	.60
	(17.54)	(9.14)	(3.48)	(2.72)	(15.92)	(10.87)	(2.15)	(9.84)	(4.36)	(5.90)	(5.26)	(4.07)	(5.21)
1980	.19	.69	.29	.16	.91	.75	.16	.68	.28	.40	.54	.48	.58
	(13.16)	(9.46)	(3.22)	(3.21)	(14.76)	(11.07)	(1.86)	(9.76)	(3.88)	(5.66)	(5.04)	(4.19)	(4.96)
1981	.17	.77	.27	.24	.89	.83	.14	.76	.26	.48	.52	.56	.56
	(9.61)	(10.58)	(2.88)	(4.59)	(14.39)	(12.81)	(1.63)	(10.60)	(3.16)	(6.04)	(4.86)	(4.86)	(4.82)

Figures in parenthesis are t-values

using stacked district data (Method III) and using the stacked district data by regions (Method IV), is presented in Table 8.15 and Table 8.16. Tables 8.11 to 8.14 show that, contrary to what would be expected, the division of the country into two regions has not improved the results of the estimation procedures. In general, the t-values at the 5% level of significance associated with the transition probability estimate are better when the stepwise regression is performed on all stacked district data (Method III), for both districts in the analysis. However, as already noted both OLS estimation procedures, Method III and Method IV, have given distorted estimates for the transition probabilities corresponding to terminal or first grades of levels of education.

Tables 8.15 and 8.16 show that the unrestricted OLS stepwise regression using Method III is the estimation procedure that generates better correlation coefficients when comparing the OLS estimates of the transition probabilities with the corresponding observed point estimates. This estimation procedure is the one that gives better Durbin-Watson statistics showing no evidence for the existence of serial correlation, with the exception, however, of Braganca, where the observed and estimated transition probabilities $p_{10,11}$ are serially correlated.

Before going further in the estimation procedures, and because the division of the country into regions has not produced more reliable transition probability estimates, a new attempt was made for each region, performing first the stepwise regression on the supply-side and on the demand-side explanatory variables separately, and using afterwards only the selected significant explanatory variables in a new OLS stepwise regression method (Method V). The

TABLE 8.15: **Comparison between the Estimates Obtained for the Stepwise Regression on Stacked District Data and the Stepwise Regression for Region 1 - LISBOA**

Transition Probability	Observed Point Estimate		Method III Stepwise Regression on Stacked District Data (All Districts)							
			Unrestricted OLS				Restricted OLS			
	Mean Value	Stand Dev.	Mean Value	Stand Dev.	R	D-W	Mean Value	Stand Dev.	R	D-W
$p_{5,5}$	.1655	.0375	.1602	.0402	.87727	1.92	.1398	.0558	.78402	2.37
$p_{5,6}$	.7673	.0544	.6901	.0876	.80433	1.68	.7301	.0946	.77193	1.63
$p_{6,6}$	.1709	.0459	.3002	.0402	.65430	2.32	.2398	.0558	.53666	2.89
$p_{6,7}$	.8255	.1356	.3501	.0876	.83211	2.57	.2000	.0946	.85607	2.72
$p_{7,7}$	.1964	.0524	.7002	.0402	.71054	1.07	.8598	.0558	.54462	.63*
$p_{7,8}$	.6982	.1292	.6901	.0876	.65036	1.63	.7901	.0946	.67369	1.68
$p_{8,8}$	.2000	.0422	.2602	.0402	.76712	1.08	.1098	.0558	.66650	.68*
$p_{8,9}$	.6900	.1039	.8201	.0876	.67890	1.03	.7201	.0946	.62458	1.30
$p_{9,9}$	.2282	.0623	.1302	.0402	.56709	2.06	.2298	.0558	.75321	1.78
$p_{9,10}$	.6909	.1437	.3301	.0876	.86383	1.24	.4401	.0946	.87515	1.61
$p_{10,10}$	.0836	.0338	.7002	.0402	-.09822	1.98	.4898	.0558	-.00014	2.09
$p_{10,11}$	.8482	.1227	.7301	.0876	.22825	1.85	.5201	.0946	.21318	1.97
$p_{11,11}$	.1836	.0717	.3602	.0402	.46238	1.04	.5298	.0558	.48938	1.47

TABLE 8.15: Comparison between the Estimates Obtained for the Stepwise Regression on Stacked District Data and the Stepwise Regression for Region 1 - LISBOA (Continued)

Transition Probability	Method IV - Stepwise Regression (Region 1)							
	Unrestricted OLS				Restricted OLS			
	Mean Value	Stand Dev.	R	D-W	Mean Value	Stand Dev.	R	D-W
$P_{5,5}$	.1600	.0383	.58389	1.44	.1499	.0261	.05666	1.53
$P_{5,6}$	.7698	.0725	.70171	1.45	.8400	.0790	.71402	1.51
$P_{6,6}$	.2100	.0383	.31493	2.34	.1099	.0261	.06924	2.51
$P_{6,7}$	.4498	.0725	.76821	2.16	.2600	.0790	.77113	2.17
$P_{7,7}$	.6000	.0383	.52715	.62*	.7899	.0261	-.27779	.45*
$P_{7,8}$	.7298	.0725	.69187	1.68	.8100	.0790	.72519	1.36
$P_{8,8}$	.2100	.0383	.48018	.89*	.0899	.0261	-.20136	.98*
$P_{8,9}$	.8198	.0725	.61810	1.21	.6600	.0790	.71678	.89*
$P_{9,9}$	.1300	.0383	.32162	2.11	.2899	.0261	.18551	2.35
$P_{9,10}$	.3598	.0725	.75936	1.40	.5300	.0790	.85610	1.08
$P_{10,10}$	.6700	.0383	.23218	2.42	.3399	.0261	.36223	2.30
$P_{10,11}$	.7498	.0725	.20374	1.85	.5800	.0790	.21621	1.88
$P_{11,11}$	.3400	.0383	.39562	1.20	.4699	.0261	.63396	2.13

TABLE: 8.16: Comparison between the Estimates Obtained for the Stepwise Regression on Stacked District Data and the Stepwise Regression Performed on Region 2 - BRAGANCA

Transition Probability	Observed Point Estimate		Method III - Stepwise Regression on Stacked District Data (All Districts)							
			Unrestricted OLS				Restricted OLS			
	Mean Value	Stand. Dev.	Mean Value	Stand. Dev.	R	D-W	Mean Value	Stand. Dev.	R	D-W
$P_{5,5}$	.1391	.0519	.1602	.0453	.87291	1.48	.1399	.0478	.53162	.71*
$P_{5,6}$	.7445	.0667	.6900	.0596	.58398	2.11	.7299	.0524	.45272	1.58
$P_{6,6}$	.1555	.0530	.3002	.0453	.82007	1.16	.2399	.0478	.61918	1.25
$P_{6,7}$	.6627	.1026	.3500	.0596	.44161	2.54	.1999	.0524	.33819	1.96
$P_{7,7}$	.2427	.1070	.7002	.0453	.71578	2.13	.8599	.0478	.43268	2.12
$P_{7,8}$	.6555	.1349	.6900	.0596	.28482	1.40	.7899	.0524	.47600	1.39
$P_{8,8}$	.2455	.0928	.2602	.0453	.66648	1.97	.1099	.0478	.48271	1.79
$P_{8,9}$	.7536	.2086	.8200	.0596	.87806	1.90	.7199	.0524	.64209	1.39
$P_{9,9}$	.2845	.1055	.1202	.0453	.60902	1.61	.2299	.0478	.76118	1.55
$P_{9,10}$	.4965	.1782	.3300	.0596	.84316	2.39	.4399	.0524	.74925	1.16
$P_{10,10}$	.0336	.0408	.7002	.0453	.65987	2.93	.4899	.0478	.66602	2.55
$P_{10,11}$	.7627	.1592	.7300	.0596	-.04022	.83*	.5199	.0524	.02292	.83*
$P_{11,11}$	.2200	.1780	.3602	.0453	.49641	2.10	.5299	.0478	.25041	2.31

TABLE: 8.16: Comparison between the Estimates Obtained for the Stepwise Regression on Stacked District Data and the Stepwise Regression Performed on Region 2 - **BRAGANCA** (Continued)

Transition Probability	Method IV - Stepwise Regression (Region 2							
	Unrestricted OLS				Restricted OLS			
	Mean Value	Stand. Dev.	R	D-W	Mean Value	Stand. Dev.	R	D-W
$P_{5,5}$	.1600	.0305	.54570	.59*	.1184	.0385	.84259	1.37
$P_{5,6}$	.3202	.0739	.17711	1.18	.4161	.0559	.81364	1.68
$P_{6,6}$	.6900	.0305	.63547	1.51	.5384	.0385	.47384	2.20
$P_{6,7}$	.2502	.0739	.03382	1.39	.2161	.0559	.84755	2.50
$P_{7,7}$	.7800	.0305	.73089	2.14	.7484	.0385	.47894	2.69
$P_{7,8}$	.5402	.0739	.34367	1.32	.6661	.0559	.53851	1.64
$P_{8,8}$	.3900	.0305	.49919	1.11	.2484	.0385	.54211	2.39
$P_{8,9}$	.7302	.0739	.68255	.95*	.9661	.0559	.54853	2.81
$P_{9,9}$	.2100	.0305	.10122	1.49	.0284	.0385	.45032	1.93
$P_{9,10}$	.2402	.0739	.62161	1.27	.1361	.0559	.61432	1.70
$P_{10,10}$	.7300	.0305	-.01364	1.60	.8384	.0385	.73833	3.12
$P_{10,11}$	.8002	.0739	-.00413	.83*	.8861	.0559	-.08258	.94*
$P_{11,11}$	.2900	.0305	.02947	2.12	.2084	.0385	.36046	2.38

estimates of the $\underline{a}$, $\underline{\delta}_m$ and $\underline{\delta}_p$ parameters of the model are presented in Appendix F, Tables F.19 and F.20. The transition probability estimates and corresponding t-values are presented in Tables F.21-F.24, and the correlation coefficients and Durbin-Watson statistics resulting from the regression of the observed point estimates of the transition probabilities on the OLS stepwise regression estimates are presented in Table F.25.

The results of this attempt show that it gives $\underline{\hat{a}}$ vector estimates very similar to the $\underline{a}$ vector estimates obtained with a straightforward application of the OLS stepwise regression to each region individually. Also, the t-values do not show significant differences when both attempts were performed (Method IV and Method V). Therefore a chi-square one sample test was used to discover whether the t-value series are significantly different from each other. The result of the test confirmed the existence of no significant differences in the t-values of the two estimated series for each transition probability.

Table F.25 in Appendix F shows that, in general, this attempt has given better correlation coefficients than the previous Method IV for Region 1, either using the unrestricted OLS stepwise regression or applying the restricted OLS stepwise regression. However, the same cannot be concluded for Region 2, where the corresponding correlation coefficients are much lower, especially when the row-sum condition has not been embodied in the model. A comparison between Table F.25, Table 8.15 and Table 8.16 show that the unrestricted OLS stepwise regression using all stacked district data (Method III) is still the estimation procedure that gives less absolute values differences, no serial correlations of the residuals, and better correlation coefficients, which means transition probabilities estimate patterns more similar to the corresponding observed point estimates. Nevertheless, it must be noted that for the district of Lisboa the correlation coefficient associated with the transition probabilities related with the last grades of the school subsystem in analysis, are higher when the Region 1 is eliminated from the stepwise regression; the corresponding transition probability estimates are, however, very far from the observed point estimates.

Finally, it is interesting to note that variable HELP and variable UNEMP have been selected to determine changes in the promotion probabilities in all the unrestricted and restricted OLS stepwise estimation procedures. The analysis of the signs explaining how the exogenous variables influence the changes of the transition probabilities will be examined in the next chapter.

8.4. The Pooled Cross-Section Time-Series Estimator

In order to obtain efficient estimators of the transition probabilities, the generalised least squares (GLS) estimation procedure must be used. The assumption of normally distributed disturbance terms with constant variances over the observations is usually not fulfilled by the OLS estimator. Heteroscedasticity is present, though the estimated variances of the estimated transition probabilities are not the minimum variances, these variances also being biased estimators of the true variances of the estimated probabilities.

A significant difference exists, however, between the basic Markov model and the extended Markov model applied to the education system. While the basic Markov model can be described by a set of equations, each equation determining the number of students in a certain grade of the school system, the extended Markov model is described by one equation only. This is due to the assumption that the explanatory variables affect in the same way all the repetition probabilities and all the promotion probabilities, respectively. Therefore, different GLS estimation procedures have been applied to each model. The method proposed by ZELLNER [1962] for estimating seemingly unrelated regressions and described in Chapter 4, was used for the basic Markov model. As the extended Markov model is described by one equation only, ZELLNER's method cannot be applied in this case, and another method must be used to estimate the variance-covariance matrix of the disturbance terms. Therefore, the pooled cross-section time-series estimation procedure described by KMENTA [1971, pp509-512] was used. This method presupposes a certain number of cross-sectional units (in this study, districts) and combines the assumptions frequently made about cross-sectional observations with those that are usually made when dealing with time-series. For the cross-sectional observations, that is, for the district observations at one year, the method assumes that the disturbance terms are mutually independent but heteroscedastic. Concerning the time-series data, the method assumes that the disturbances are autoregressive, though not necessarily heteroscedastic. Thus, heteroscedasticity, cross-sectional independence[1] and autoregression are the characteristics assumed for the GLS estimation procedure applied to the extended Markov model.

The pooled cross-section time-series method assumes, then, that the variance-covariance matrix of the disturbance terms, Ω, has the following form

$$\underline{\Theta} = \begin{bmatrix} \sigma_1^2 \, \underline{\Gamma} & \underline{0} & \cdots & \underline{0} \\ \underline{0} & \sigma_2^2 \, \underline{\Gamma}_2 & \cdots & \underline{0} \\ \cdots & \cdots & \cdots & \cdots \\ \underline{0} & \underline{0} & \cdots & \sigma_N^2 \, \Gamma_N \end{bmatrix}$$

where

$$\underline{\Gamma_i} = \begin{bmatrix} 1 & \underline{\rho_i} & \underline{\rho_i^2} & \cdots & \underline{\rho_i}^{sk-1} \\ \underline{\rho_i} & 1 & \underline{\rho_i} & \cdots & \underline{\rho_i}^{sk-2} \\ \cdot\cdot & \cdot\cdot & \cdot\cdot & \cdots & \cdot\cdot \\ \underline{\rho_i}^{sk-1} & \underline{\rho_i}^{sk-2} & \underline{\rho_i}^{sk-3} & \cdots & 1 \end{bmatrix}$$

with each $\underline{0}$ matrix being a (sk x sk) matrix of zeros, sk being the number of observations per cross-section (grades x time) and N being the number of cross-sections of the data set.

The consistent estimator of ρ_i is given by [see KMENTA (1971),pp.510-511]

$$\underline{\hat{\rho}_i} = \sum_t w_i^t \, w_i^{t-1} \;\; / \sum_t (w_i^{t-1})^2 \qquad t=2, \; \ldots \; ,2k$$

and the consistent estimator of σ_i can be obtained by using $\hat{\rho}_i$ to transform the original data and using the OLS method with the transformed data. The variances of the residuals of this estimation procedure are estimated by

$$s_{wi}^2 = [\; 1 \,/(sk - q - 1)] \sum_{t=2}^{2k} (w_i^t)^2$$

where q is the number of independent variables of the model (in the study q = 2s + 2m - 1). Thus, the estimator of σ_i can be obtained using

$$s_i^2 = s_{wi}^2 / (1 - \rho_i^2)$$

For the present study the variance-covariance matrix Θ has size (1386 x 1386) for the unrestricted GLS estimation procedure and size (1584 x 1584) for the restricted GLS estimation procedure. Pooled cross-section time-series estimation procedures were attempted for all the stepwise OLS regressions previously performed in this chapter, using the SHAZAM econometric computer program with the Pool regression subproblem command (see Appendix F, program REGSXA3A, for Region 1; a similar program has been used to estimate the coefficients for Region 2). However, because of an insufficient amount of memory for internal SHAZAM workspace, the pooled cross-section time-series estimation procedure has only been run for the unrestricted OLS (Method IV) and for the unrestricted OLS (Method V), that is, when the Θ matrix has the smaller size (693 x 693). Thus the pooled cross-section time-series estimation regression has only been performed by regions, either after using all the explanatory variables in the stepwise regression (Method VI) or after using only the significant explanatory variables after performing stepwise regression on the supply-side and on the demand-side explanatory variables separately (Method VII). The estimated values of the coefficients a, δ_r and δ_p corresponding to Method VI are presented in Table 8.17 and the transition probability estimates are presented in Table 8.18 and Table 8.19. The results obtained after performing the pooled cross-section time-series estimation procedure with the selected explanatory variables obtained by Method V are presented in Appendix F, Tables F.26-F.28.

TABLE 8.17: **Pooled Cross-Section Time-Series Estimation Using the Significant Variables of the Unrestricted OLS with All Explanatory Variables (Method VI)**

Region 1				Region 2			
Coefficient	Estimated Value	St. Error	t-value	Coefficient	Estimated Value	St. Error	t-value
a_{55}	.16	.0118	13.54	a_{55}	.14	.0082	17.45
a_{56}	.47	.0896	5.24	a_{56}	.31	.0801	3.87
a_{66}	.51	.1084	4.66	a_{66}	.68	.0980	6.98
a_{67}	.40	.0669	6.00	a_{67}	.17	.0455	3.79
a_{77}	.58	.0785	7.37	a_{77}	.84	.0585	14.30
a_{78}	.69	.0601	11.43	a_{78}	.53	.0666	7.88
a_{88}	.22	.0706	2.96	a_{88}	.41	.0809	5.11
a_{89}	.88	.0678	12.92	a_{89}	.72	.0715	10.08
a_{99}	.53	.0769	.69*	a_{99}	.26	.0792	3.27
$a_{9,10}$	.26	.0709	3.67	$a_{9,10}$	.28	.0635	4.37
$a_{10,10}$	.67	.1186	5.61	$a_{10,10}$	.62	.1095	5.66
$a_{10,11}$	.68	.1252	5.43	$a_{10,11}$	.78	.1119	6.97
$a_{11,11}$	.40	.1313	3.01	$a_{11,11}$	.31	.1184	2.60
EDUC 1	.014	.0051	2.73	TEARN 1	.009	.0089	.95*
GDP 1	.037	.0092	4.02	BUS 1	-.010	.0625	.79*
PCAP 2	-.002	.0075	2.17	HELP 1	.002	.0135	.13*
TEARN 2	-.018	.0053	3.34	GDP 1	0030	.0103	2.87
HELP 2	.042	.0082	5.13	EARN 1	.010	.0094	1.04*
UNEMP 2	-.049	.0113	4.44	PCAP 2	-.037	.0092	4.05
LFLEV 2	-.047	.0118	3.94	TEARN 2	.003	.0088	.31*
				HELP 2	.033	.0128	2.55
				UNQUAL 2	.035	.0134	2.60
				UNEMP 2	-.063	.0117	5.37

R^2 = .9675 R^2 = .9856

$\bar{R}^2$ = .9666 $\bar{R}^2$ = .9851

df = 672 df = 669

dw = 2.1881 dw = 2.0604

$t^*_{0.025}$ = 1.96

TABLE 8.18: Method VI - Transition Probability Estimates Using Pooled Cross-section Time-series Regression Using the Significant Variables of Method IV - Region 1 - LISBOA

YEAR	P_{11}	P_{12}	P_{22}	P_{23}	P_{33}	P_{34}	P_{44}	P_{45}	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}
1971	.13	.49	.51	.44	.57	.74	.18	.93	.03	.31	.66	.73	.36
	(5.65)	(5.15)	(4.37)	(5.57)	(6.63)	(10.10)	(2.14)	(11.09)	(.35)	(3.32)	(4.90)	(5.15)	(2.58)
1972	.14	.46	.52	.41	.58	.71	.19	.90	.04	.28	.67	.70	.37
	(6.79)	(4.85)	(4.49)	(5.63)	(6.86)	(10.03)	(2.28)	(11.85)	(.46)	(3.26)	(5.10)	(5.22)	(2.68)
1973	.15	.49	.53	.44	.59	.74	.20	.93	.05	.31	.68	.73	.38
	(8.17)	(5.17)	(4.60)	(6.25)	(7.07)	(10.71)	(2.44)	(12.51)	(.58)	(3.41)	(5.29)	(5.42)	(2.78)
1974	.08	.58	.46	.53	.52	.83	.13	1.02	-.02	.40	.61	.82	.31
	(2.97)	(5.89)	(3.98)	(6.44)	(5.89)	(11.39)	(1.57)	(12.38)	(.16)	(3.93)	(4.47)	(6.00)	(2.14)
1975	.12	.48	.50	.43	.56	.73	.17	.92	.02	.30	.65	.72	.35
	(5.99)	(4.88)	(4.37)	(5.57)	(6.72)	(10.83)	(2.13)	(12.17)	(.20)	(3.30)	(5.09)	(6.02)	(2.50)
1976	.14	.48	.52	.43	.58	.73	.19	.92	.04	.30	.67	.72	.37
	(9.89)	(5.23)	(4.63)	(5.89)	(7.20)	(12.03)	(2.52)	(12.08)	(.52)	(3.76)	(5.50)	(5.50)	(2.74)
1977	.18	.42	.56	.37	.62	.67	.23	.86	.08	.24	.71	.66	.41
	(15.00)	(4.41)	(4.97)	(5.59)	(7.63)	(9.32)	(2.96)	(12.72)	(1.02)	(3.26)	(5.93)	(4.99)	(3.06)
1978	.19	.34	.57	.29	.63	.59	.24	.78	.09	.16	.72	.58	.42
	(15.11)	(3.59)	(5.04)	(4.01)	(7.67)	(8.47)	(3.03)	(9.66)	(1.12)	(2.43)	(6.02)	(4.17)	(3.10)
1979	.21	.34	.59	.29	.65	.39	.26	.78	.11	.16	.74	.58	.44
	(13.50)	(3.57)	(5.16)	(4.00)	(7.68)	(7.91)	(3.12)	(9.41)	(1.28)	(2.20)	(6.12)	(4.09)	(3.17)
1980	.22	.38	.60	.33	.66	.63	.27	.82	.12	.20	.75	.62	.45
	(12.69)	(3.95)	(5.22)	(4.70)	(7.66)	(8.50)	(3.17)	(10.37)	(1.37)	(2.47)	(6.18)	(4.47)	(3.21)
1981	.21	.39	.59	.34	.65	.64	.26	.83	.11	.21	.74	.63	.44
	(12.63)	(4.18)	(5.18)	(4.79)	(7.62)	(10.15)	(3.13)	(10.64)	(1.27)	(3.00)	(6.19)	(4.66)	(3.13)

Figures in parenthesis are t-values

TABLE 8.19: Method VI - Transition Probability Estimates Using Pooled Cross-section Time-series Regression Using The Significant Variables of Method IV - Region 2 - BRAGANCA

YEAR	P_{11}	P_{12}	P_{22}	P_{23}	P_{33}	P_{34}	P_{44}	P_{45}	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}
1971	.12 (4.57)	.34 (4.08)	.63 (5.83)	.19 (3.06)	.81 (10.95)	.55 (6.31)	.37 (3.69)	.75 (8.17)	.21 (2.05)	.28 (3.05)	.61 (4.53)	.80 (5.86)	.27 (1.96)
1972	.15 (7.20)	.34 (4.13)	.66 (6.22)	.19 (3.58)	.84 (12.17)	.55 (7.06)	.40 (4.28)	.75 (9.49)	.24 (2.56)	.28 (3.37)	.64 (5.11)	.80 (6.46)	.30 (2.24)
1973	.16 (8.90)	.34 (4.10)	.67 (6.35)	.19 (3.61)	.85 (12.64)	.55 (7.06)	.41 (4.46)	.75 (9.59)	.25 (2.71)	.28 (2.71)	.65 (5.26)	.80 (6.42)	.31 (2.33)
1974	.10 (5.37)	.40 (4.74)	.61 (5.76)	.25 (4.78)	.79 (11.78)	.61 (8.86)	.35 (4.16)	.81 (10.33)	.19 (2.05)	.34 (4.32)	.59 (4.79)	.86 (6.78)	.25 (1.88)
1975	.10 (3.41)	.42 (4.45)	.61 (5.58)	.27 (4.01)	.79 (10.19)	.63 (7.59)	.35 (3.43)	.83 (8.96)	.19 (1.89)	.36 (3.58)	.59 (4.47)	.88 (7.35)	.25 (1.90)
1976	.13 (13.10)	.35 (4.21)	.64 (6.25)	.20 (3.91)	.82 (13.30)	.56 (8.53)	.38 (4.50)	.76 (9.69)	.22 (2.69)	.29 (3.95)	.62 (5.38)	.81 (6.80)	.28 (2.32)
1977	.14 (8.59)	.30 (3.37)	.65 (6.21)	.15 (2.88)	.83 (12.37)	.51 (6.49)	.39 (4.70)	.71 (8.81)	.23 (2.66)	.24 (3.39)	.63 (5.31)	.76 (5.99)	.29 (2.23)
1978	.16 (15.13)	.24 (2.79)	.67 (6.54)	.09 (1.78)	.85 (13.50)	.45 (5.95)	.41 (4.85)	.65 (7.90)	.25 (3.12)	.18 (2.93)	.65 (5.72)	.70 (5.71)	.31 (2.54)
1979	.17 (15.08)	.26 (3.01)	.68 (6.56)	.11 (2.18)	.86 (13.49)	.47 (6.40)	.42 (4.84)	.67 (8.33)	.26 (3.14)	.20 (3.17)	.66 (5.73)	.72 (6.00)	.32 (2.57)
1980	.17 (11.77)	.28 (3.19)	.68 (6.49)	.13 (2.50)	.86 (13.06)	.49 (6.30)	.42 (4.69)	.69 (8.41)	.26 (3.00)	.22 (3.04)	.66 (5.56)	.74 (6.11)	.32 (2.52)
1981	.15 (9.38)	.35 (4.09)	.06 (6.27)	.20 (3.66)	.84 (12.64)	.56 (7.93)	.40 (4.46)	.76 (9.11)	.24 (2.69)	.29 (3.47)	.64 (5.29)	.81 (6.40)	.30 (2.34)

Figures in parenthesis are t-values

Concerning the coefficient estimates, and comparing Table 8.17 with Tables 8.5 and 8.6, it is apparent that although for Region 1 the standard errors are lower and the t-values of the the estimated coefficients are better, when performing the pooled cross-section estimation procedure for Region 2[2], the standard errors have become larger and consequently[3] the t-values are lower, some of the coefficients even showing non-significant t-values at the 20% level. However, on running the FORTRAN programs STAT(CON) and ERROR(CASE) adapted to the corresponding data to produce the transition probability estimates and associated t-values, Region 1 presents, in general, lower t-values and Region 2 presents better t-values than those obtained when performing the unrestricted stepwise OLS regression. The pooled cross-section time-series estimation procedure applied to Region 2 has generated significant t-values at the 5% level of significance for all the transition probability estimates; the method has also improved the reliability of the transition probability estimates for Region 1, presenting significant t-values for the repetition probability estimate. However, for the repetition probability estimate the t-values are still very low.

Analogous conclusions result when comparing Tables F.26 - F.28 with Tables F.19 - F.24 presented in Appendix F. It must be noted, however, that the pooled cross-section time-series estimation procedure using only the previously selected supply-side and demand-side explanatory variables (Method VII) has produced significant t-values at the 5% level of significance for all the transition probability estimates and for both regions.

Table 8.20 presents the results of performing the regression of the observed point estimates of the transition probabilities on the

TABLE 8.20: Comparison between the Estimates Obtained for the pooled cross-section time-series Estimation Procedures (Unrestricted OLS)

	LISBOA								BRAGANCA							
	Method VI - Pooled Cross-section time-series on the Results of Method IV				Method VII - Pooled Cross-section time-series on the Results of Method V				Method VI - Pooled Cross-section time-series on the Results of Method IV				Method VII - Pooled Cross-section time-series on the Results of Method V			
	Mean Value	Stand Dev.	R	D-W	Mean Value	Stand Dev.	R	D-W	Mean Value	Stand Dev.	R	D-W	Mean Value	Stand Dev.	R	D-W
P_{55}	.1599	.0375	.58688	1.42	.1601	.0422	.73199	2.16	.1401	.0229	.88572	1.45	.1399	.0240	.74862	1.04
P_{56}	.4701	.0644	.67155	1.54	.4401	.0764	.79984	1.35	.3101	.0548	.54377	1.69	.3301	.0543	.68812	1.59
P_{66}	.5099	.0375	.32195	2.33	.5401	.0422	.51487	2.67	.6801	.0229	.56808	1.99	.6499	.0240	.72063	2.12
P_{67}	.4001	.0644	.72821	2.26	.3901	.0764	.89716	2.56	.1701	.548	.59852	2.21	.1801	.0543	.69993	2.28
P_{77}	.5799	.0375	.52087	.62*	.6001	.0422	.59152	.61*	.8401	.229	.71412	2.79	.8299	.0240	.68246	2.79
P_{78}	.6901	.0644	.61723	1.90	.6901	.0764	.55707	2.06	.5301	.548	.58175	1.80	.5401	.0543	.47024	1.61
P_{88}	.2199	.0375	.47198	.89*	.2101	.0422	.57172	.99	.4101	.229	.66559	2.12	.3899	.0248	.53893	2.23
P_{89}	.8801	.0644	.53553	1.43	.8801	.0764	.50101	1.66	.7201	.548	.64344	1.54	.7401	.0543	.59907	2.06
P_{99}	.5299	.0375	.32378	2.11	.0601	.0422	.59621	2.38	.2601	.229	.37270	1.65	.2299	.0240	.53826	1.90
$P_{9,10}$	.2601	.0644	.66419	1.68	.2601	.0764	.53424	1.88	.2801	.548	.66645	1.30	.2701	.0543	.76288	1.68
$P_{10,10}$	.6699	.0375	.22616	2.41	.6901	.0422	.07248	2.14	.6201	.229	.58048	2.63	.6299	.0240	.49287	2.49
$P_{10,11}$	.6701	.0644	.08916	2.00	.6801	.0764	.20957	1.97	.7801	.548	-.02914	.85*	.7901	.0543	.11331	.76*
$P_{11,11}$	.3999	.0375	.39820	1.20	.3901	.0422	.60311	1.27	.3101	.229	.35627	2.20	.2899	.0240	.56818	1.98

corresponding pooled cross-section time-series estimates, for the two districts Lisboa and Braganca. Comparing the results obtained for both cases (Method VI and Method VII), the table shows that the correlation coefficients for Lisboa are better for ten out of the thirteen transition probabilities, using the transition probability estimates of Method VII. However, only slightly better correlation coefficients (seven out of thirteen) have occurred for the district of Braganca. These results are not unexpected because the stepwise OLS estimation procedure using the selected explanatory variables after having performed the same method on the supply-side and on the demand-side explanatory variables separately (Method V), has proved to be better than the stepwise OLS estimation procedure performed on all stacked section data and all explanatory variables (Method IV), for the district of Lisboa (Region 1). Furthermore, for the district of Braganca (Region 2), Method IV has proved to give slightly better time-patterns when the restricted OLS estimation procedure was performed.

A comparison between these correlation coefficients (Method VI and Method VII) and the Durbin-Watson statistics, and the corresponding statistics obtained after performing Method IV and Method V, reveals that Method VII did not generate better correlation coefficients and therefore no closer time-patterns were obtained for the district of Lisboa when the unrestricted GLS estimation procedure was performed. It has, however, generated appreciably better correlation coefficients for Braganca. Nevertheless, these better correlation coefficients are not sufficiently good to suggest that the transition probability estimates for Braganca describe in a satisfactory way the time-patterns of the observed point estimates of the transition probabilities.

The unrestricted stepwise OLS estimation procedure using all stacked district data (Method III) has proved to be the OLS method; this gives least absolute values bias, better time-patterns and no serially correlated residuals when the regression of the observed point estimates on the transition probability estimates is performed. It seems, then, worthwhile to compare the results of this method with the results obtained after performing the pooled cross-section time-series estimation (Method VII). The comparison shows that although the transition probability estimates are not reasonable for the $\hat{p}_{67}$, $\hat{p}_{77}$, $\hat{p}_{9,10}$ and $\hat{p}_{10,10}$ estimates, Method III still presents closer mean values for the transition probability estimates and, in general, similar time-patterns. It must be noted, however, that very low correlation coefficients have occured for the transition probability estimates $\hat{p}_{10,11}$ and for all methods performed in this study. Although the correlation coefficients are, in general, better for Lisboa than for Braganca, when the unrestricted OLS estimation procedure using all stacked district data (Method III) is performed, the transition probability estimates $\hat{p}_{10,10}$ also have very low correlation coefficients for that district, thus different time-patterns are observed between its observed point estimate and the corresponding transition probability estimates.

8.5 Principal Components Analysis

8.5.1 Principal Components Analysis Applied to the Different Districts

Following the procedure in the previous chapter for the whole country, principal components analysis is now performed by district in order to reduce the degree of multicollinearity between the selected explanatory variables. The method used for the whole country is identically applied to each district, individually.

Tables F.29 - F.32 presented in Appendix F, describe the factor matrices and factor loadings obtained after using the subprogram FACTOR included in the SPSS package. As expected, the tables show that the higher percentage variance values occur in the district where a second (for the demand-side variables) or a third (for the supply-side variables) principal component is selected. All districts present values over 77% for the percentage variance, which means that in all districts the selected principal components can be used as good proxies of the set of the original explanatory variables.

The standardised values of the principal components are also presented in Appendix F, Table F.33 and Table F.34. The patterns of the factor loadings for most of the districts suggest interpretations analogous to the interpretations formulated for the whole country. Ten districts present only two principal components for the supply-side variables, which can be identified in the same way as Factor I and Factor II were described for the whole country. However, in the case of Portalegre, Porto and Santarem, Factor II includes a significant loading for variable PUCLASS, which indicates that in these districts Factor II is not only representative of the teachers motivation, but also representative of the school places offered to the students. The remaining districts show three principal components identical to Factors I, II and III observed for the whole country. It is important, however, to note that while Factor III can be identified for the whole country as an inverse measure of the school places offered to the students, the principal components analysis applied to each district shows, with the exception for Braganca, that Factor III is now interpreted as a direct measure of the school places offered to the students[4].

The principal components analysis of the explanatory variables of the demand-side of the education system shows that eight districts present significant mean earnings to create a new principal component. Therefore, although Factor I still represents a general index of the well-being of the population, Factor II may reflect expectations, as it broadly represents the salaries level of the population.

8.5.2 The OLS Estimator with Principal Components

In the previous section, different sets of principal components were obtained for the eighteen districts separately, but for example, if three principal components have been found, they can correspond to two supply-side components and one demand-side component, or vice-versa. The assumption of equal influence of the explanatory variables (now, the principal components) on the changes of the repetition probabilities for all districts or on the changes of the promotion probabilities for all districts, is then not reasonable when principal components are used. Thus, in this section different vectors $\underline{\delta}_r$ and $\underline{\delta}_p$ are estimated by district, the $\underline{a}$ coefficients, however, being the same for all districts (Method VIII). Furthermore, district dummy variables have been included in the model. Seventeen dummy variables have been constructed to account for the district differences. Lisboa was the district chosen to have zero values. The estimated coefficient of a dummy variable represents, therefore, the difference associated with a change from Lisboa to the corresponding district. Thus, matrix N^x now has size (1386 x 170) for the unrestricted OLS estimation procedure and (1584 x 170) for the restricted OLS estimation procedure. The vector of the $\underline{a}$ coefficients has, as usual size (13 x 1), the $\underline{\delta}$ vector has size (140 x 1) and the dummy variables vector has size (17 x 1).

District data files have been generated using the same process undertaken for the whole country when principal components were used. Concerning the unrestricted OLS estimation procedure, these files have been stacked in file STADATA and FORTRAN program PROGRM(STACK) has been used to generate the data file matrix $[\underline{n}^* \vdots \underline{N}^x]$, which has the following form

$$
\begin{bmatrix}
\underline{n}_1^* & \underline{N}_1^o & \underline{n}_1^{t-1} \otimes \underline{x}_1 & \underline{0} & \cdots & \underline{0} & \underline{1} & \underline{0} & \underline{0} & \underline{0} & \cdots & \underline{0} \\
\underline{n}_2^* & \underline{N}_2^o & \underline{0} & \underline{n}_2^{t-1} \otimes \underline{x}_2 & \cdots & \underline{0} & \underline{0} & \underline{1} & \underline{0} & \underline{0} & \cdots & \underline{0} \\
\underline{n}_3^* & \underline{N}_3^o & \underline{0} & \underline{0} & \cdots & \underline{0} & \underline{0} & \underline{0} & \underline{1} & \underline{0} & \cdots & \underline{0} \\
\underline{n}_{12}^* & \cdot & \cdot & \cdot & \cdots & \cdot & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \cdots & \underline{0} \\
\underline{n}_{18}^* & \cdot & \cdot & \cdot & \cdots & \cdot & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \cdots & \underline{1} \\
\underline{n}_4^* & \cdot & \cdot & \cdot & & \cdot & \underline{0} & \underline{0} & \underline{0} & \underline{1} & \cdots & \underline{0} \\
\underline{n}_6^* & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\
\underline{n}_7^* & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\
\underline{n}_8^* & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\
\underline{n}_9^* & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\
\underline{n}_{11}^* & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\
\underline{n}_{13}^* & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\
\underline{n}_{14}^* & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\
\underline{n}_{16}^* & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\
\underline{n}_{17}^* & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\
\underline{n}_5^* & \cdot & \cdot & \cdot & & \cdot & \cdot & \cdot & \cdot & \cdot & & \cdot \\
\underline{n}_{10}^* & \underline{N}_{10}^o & \underline{0} & \underline{0} & \cdots & \cdot & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \cdots & \underline{0} \\
\underline{n}_{15}^* & \underline{N}_{15}^o & \underline{0} & \underline{0} & \cdots & \underline{n}_{15}^{t-1} \otimes \underline{x}_{15} & \underline{0} & \underline{0} & \underline{0} & \underline{0} & \cdots & \underline{0}
\end{bmatrix}
$$

The large number of resultant explanatory variables (170) makes impossible the use of SPSS or SHAZAM programs to produce the OLS estimates of the coefficients. Therefore FORTRAN program STACK was written in order to obtain these estimates, using the facilities of NAG library. FORTRAN program UNRESID was afterwards applied to produce the statistics for the coefficients. Similar programs have been used to perform the restricted OLS estimation procedure. The different estimates of the coefficients are presented in Tables 8.21 to 8.23 and the corresponding t-values are presented in Table 8.24 to Table 8.26.

Although the unrestricted and the restricted OLS estimation procedures show significant t-values at the 5% level of significance for most of the $\underline{a}$ coefficients, these estimation procedures have produced estimates for the $\underline{\hat{\delta}}$ coefficients with very low associated t-values for most of the districts; thus, even at the 20% level of significance, a large number of non-significant t-values were found for the unrestricted OLS $\underline{\hat{\delta}}$ estimates; the districts of Region 2 having, in general, lower t-values than the districts of Region 1. Moreover, even Lisboa and Porto, the districts which show higher t-values, do not present all t-values significant at the 5% level.

FORTRAN programs PROBSTAK(<name of district>), COVAR2 and ERROR(CASE) have been used to compute the transition probability estimates and corresponding t-values. The results for Lisboa and Braganca are presented in Tables 8.27 to 8.30. The inclusion of the row-sum condition into the model without imposing bounds to the transition probability estimates has caused the $\hat{p}_{89}$ estimates to be significantly greater than unity and the $\hat{p}_{99}$ estimates are significantly less than zero for both districts. The tables also

TABLE 8.21: Estimated Values of the Coefficients for the OLS Estimation Procedure with Principal Components (Method VIII)

Coefficient	Unrestricted OLS		95% Conf. Interval		Restricted OLS		95% Conf. Interval	
	Estimated Value	t-value	L.B.	U.B.	Estimated Value	t-value	L.B.	U.B.
a_{55}	.13	20.30	0.121	0.147	.15	17.10	0.129	0.162
a_{56}	.61	9.16	0.476	0.735	.66	8.75	0.510	0.803
a_{66}	.37	4.66	0.216	0.531	.32	3.51	0.143	0.504
a_{67}	.36	7.77	0.267	0.447	.32	6.34	0.223	0.423
a_{77}	.66	12.56	0.557	0.764	.71	12.09	0.599	0.830
a_{78}	.68	14.60	0.588	0.771	.76	14.31	0.657	0.865
a_{88}	.23	4.07	0.120	0.342	.15	2.29	0.021	0.277
a_{89}	.92	19.27	0.828	1.015	1.21	20.46	1.099	1.331
a_{99}	-.03	.62*	-0.142	0.074	-.36	5.26	-0.496	-0.227
$a_{9,10}$	.15	2.79	0.045	0.258	.12	1.88*	-0.005	0.252
$a_{10,10}$	.90	10.29	0.727	1.069	.95	8.97	0.745	1.161
$a_{10,11}$	.70	8.44	0.540	0.868	.83	7.71	0.618	1.039
$a_{11,11}$	.31	3.50	0.134	0.476	.19	1.65*	-0.034	0.406

Unrestricted OLS: R^2 = .97678, F = 300.62, df = 1215

Restricted OLS: R^2 = .99316, F = 1036.98, df = 1413

TABLE 8.22: Estimated Values of Delta Coefficients for the Unrestricted OLS Estimation Procedure With Principal Components

Coeff	Aveiro	Beja	Braga	Braganca	C.Branco	Coimbra	Evora	Faro	Guarda
δ_{r1}	.03254	.07842	-.08969	.06804	.04469	.02873	.00672	-.00427	-.03885
δ_{r2}	-.00017	.07439	.00218	-.01898	.02780	.01384	-.00316	.01083	.01770
δ_{r3}	-.00013	.02818	.13860	-.03378	-.00659	.04179	.00459	.05345	-.01890
δ_{r4}	-	-	-	-.04429	.00707	-.02318	.06223	-.02641	.07642
δ_{r5}	-	-	-	-	-.04494	-	-	-	-
δ_{p1}	-.07601	.03051	-.02598	-.07235	-.04054	-.05371	-.02402	.01967	-.00705
δ_{p2}	.01705	.04386	.02107	.05872	-.01511	-.01354	-.00538	.00447	.01269
δ_{p3}	.00558	-.06154	-.06090	.00666	.00202	-.04948	.01061	-.08728	.01368
δ_{p4}	-	-	-	.03162	-.01999	.04354	.05619	.01258	-.07072
δ_{p5}	-	-	-	-	.04852	-	-	-	-

TABLE 8.22: Estimated Values of the Delta Coefficients for the Unrestricted OLS Estimation Procedure with Principal Components (Continued)

Coeff	Leiria	Lisboa	Portalegre	Porto	Santarem	Setubal	V.Castelo	V. Real	Viseu
δ_{r1}	.02734	.08795	.00996	.06021	.02927	.00813	.00217	.01380	-.09483
δ_{r2}	.00369	.01667	.00171	.03359	-.02284	.00126	-.01408	.00476	-.10043
δ_{r3}	.00224	-.08104	.02520	-.01032	.02752	.02666	.03337	-.04739	.05155
δ_{r4}	.01846	.00957	-	-.05375	-.02416	.03826	-.03201	.00151	-
δ_{r5}	-.02826	-	-	-	-	-.03139	-	-	-
δ_{p1}	-.07050	-.12930	-.02746	-.05349	-.13709	.00893	.01694	.01112	.20732
δ_{p2}	-.00188	-.04478	.02828	-.01890	.12450	-.02024	.02154	.04343	.21029
δ_{p3}	-.00579	.13560	-.02627	.03707	.00267	-.04718	-.04404	.00226	-.07686
δ_{p4}	-.01479	-.02414	-	.03883	.03913	-.06447	.00623	-.03878	-
δ_{p5}	.02703	-	-	-	-	.08902	-	-	-

TABLE 8.23: Estimated Values of the Delta Coefficients for the Restricted OLS Estimation Procedure With Principal Components

Coeff	Aveiro	Beja	Braga	Braganca	C.Branco	Coimbra	Evora	Faro	Guarda
δ_{r1}	.01135	.04831	-.13102	..06760	.06024	.02128	.00382	-.00308	.01865
δ_{r2}	-.00387	.04663	.00192	-.01610	.02964	-.00037	-.01318	.01045	.00848
δ_{r3}	-.00627	.02592	.16103	-.03287	-.01396	.04632	.00547	.05323	-.02170
δ_{r4}	-	-	-	-.03703	-.00801	-.02580	.07151	-.04309	.07534
δ_{r5}	-	-	-	-	-.06007	-	-	-	-
δ_{p1}	-.06205	.01818	-.03295	-.07284	-.04259	-.06473	-.01326	.02378	.02171
δ_{p2}	.01031	.03569	.01682	.05475	-.02046	-.02782	-.01634	-.00062	.01261
δ_{p3}	.00239	-.05100	-.05059	.01159	.00306	-.03404	-.02113	-.08837	.01921
δ_{p4}	-	-	-	.04090	-.00784	.05240	-.0726	.02751	-.08809
δ_{p5}	-	-	-	-	.04970	-	-	-	-

TABLE 8.23: Estimated Values of the Delta Coefficients for the Restricted OLS Estimation Procedure With Principal Components (Continued)

Coeff	Leiria	Lisboa	Portalegre	Porto	Santarem	Setubal	V.Castelo	V.Real	Viseu
δ_{r1}	.01865	.13517	.00023	.07651	.01777	.00262	.04462	-.03780	-.09428
δ_{r2}	.00610	.01600	.00228	.01493	-.00423	-.01465	-.04726	-.00631	-.09725
δ_{r3}	.01617	-.12956	.02523	-.01101	.02758	.02238	.00989	-.04948	.03860
δ_{r4}	.01522	-.00368	-	-.02396	-.05021	.04907	-.05106	.03420	-
δ_{r5}	-.03660	-	-	-	-	-.01331	-	-	-
δ_{p1}	-.03504	-.16829	.00345	-.01638	-.12628	.02105	.08014	.01480	.23112
δ_{p2}	-.01976	-.04920	.00472	-.02699	.12468	.00576	-.03734	.03125	.23272
δ_{p3}	-.00157	.17652	-.03398	.03922	.01180	-.05495	-.06000	.02771	-.07616
δ_{p4}	-.03472	-.00778	-	.03944	.02831	-.07271	-.00499	-.02914	-
δ_{p5}	.00544	-	-	-	-	.07375	-	-	-

TABLE 8.24: t-values Associated with the Unrestricted OLS $\underline{\delta}$ Coefficients Estimates

	Aveiro	Beja	Braga	Braganca	C.Branco	Coimbra	Evora	Faro	Guarda	Leiria	Lisboa	Portalegre	Porto	Santarem	Setubal	V.Castelo	V.Real	Viseu
δ_{r1}	1.11	0.14	0.76	0.31	0.29	0.34	0.06	0.05	0.17	0.13	1.93	0.03	1.96	0.32	0.16	0.01	0.09	0.35
δ_{r2}	0.01	0.14	0.07	0.18	0.31	0.24	0.03	0.16	0.19	0.06	1.35	0.01	2.02	0.26	0.02	0.06	0.06	0.37
δ_{r3}	0.01	0.20	1.17	0.37	0.07	0.71	0.04	0.72	0.17	0.03	2.03	0.10	0.48	0.47	0.84	0.38	0.63	1.02
δ_{r4}	-	-	-	0.20	0.05	0.40	0.48	0.33	0.40	0.11	0.62	-	2.71	0.42	0.81	0.40	0.01	-
δ_{r5}	-	-	-	-	0.43	-	-	-	-	0.19	-	-	-	-	0.70	-	-	-
δ_{p1}	2.52	0.06	0.22	0.33	0.26	0.63	0.22	0.21	0.03	0.34	2.81	0.08	1.72	1.49	0.17	0.07	0.08	0.77
δ_{p2}	0.56	0.08	0.70	0.54	0.17	0.23	0.05	0.07	0.13	0.03	3.59	0.10	1.12	1.42	0.30	0.10	0.59	0.77
δ_{p3}	0.18	0.43	0.50	0.07	0.02	0.83	0.09	1.16	0.12	0.08	3.36	0.10	1.70	0.05	1.48	0.49	0.03	1.51
δ_{p4}	-	-	-	0.14	0.15	0.73	0.43	0.16	0.36	0.09	1.53	-	1.94	0.67	1.35	0.08	0.27	-
δ_{p5}	-	-	-	-	0.46	-	-	-	-	0.18	-	-	-	-	1.96	-	-	-

TABLE 8.25: t-values Associated with the Restricted OLS $\underline{\delta}$ Coefficients Estimates.

	Aveiro	Beja	Braga	Braganca	C.Branco	Coimbra	Evora	Faro	Guarda	Leiria	Lisboa	Portalegre	Porto	Santarem	Setubal	V.Castelo	V.Real	Viseu
δ_{r1}	0.28	0.06	0.81	0.23	0.29	0.18	0.03	0.02	0.11	0.06	2.12	0.00	1.80	0.14	0.04	0.14	0.19	0.25
δ_{r2}	0.09	0.06	0.05	0.11	0.24	0.01	0.08	0.11	0.07	0.07	0.92	0.01	0.65	0.04	0.16	0.15	0.06	0.26
δ_{r3}	0.15	0.13	1.00	0.26	0.11	0.56	0.03	0.51	0.14	0.17	2.32	0.07	0.37	0.34	0.50	0.08	0.47	0.56
δ_{r4}	-	-	-	0.12	-.04	0.32	0.39	0.39	0.29	0.07	0.17	-	0.88	0.62	0.75	0.47	0.18	-
δ_{r5}	-	-	-	-	0.42	-	-	-	-	0.17	-	-	-	-	0.21	-	-	-
δ_{p1}	1.49	0.02	0.20	0.23	0.19	0.52	0.09	0.17	0.07	0.12	2.52	0.01	1.37	0.94	0.28	0.24	0.07	0.60
δ_{p2}	0.24	0.05	0.39	0.35	0.16	0.33	0.10	0.01	0.09	0.23	2.70	0.01	1.11	0.98	0.06	0.12	0.29	0.60
δ_{p3}	0.05	0.25	0.30	0.09	0.02	0.39	0.12	0.80	0.12	0.02	3.02	0.09	1.25	0.14	1.18	0.48	0.26	1.05
δ_{p4}	-	-	-	0.12	0.04	0.61	0.38	0.23	0.31	0.15	0.34	-	1.37	0.34	1.05	0.04	0.14	-
δ_{p5}	-	-	-	-	0.33	-	-	-	-	0.03	-	-	-	-	1.12	-	-	-

TABLE 8.26: Estimated Coefficients and t-values for the District Dummy Variables

	Unrestricted OLS		Restricted OLS	
	Coefficient	t-value	Coefficient	t-value
Aveiro	-223.67	2.27	0.92	0.01
Beja	- 55.54	0.60	3.31	0.03
Braga	-299.63	2.99	2.67	0.02
Braganca	- 6.61	0.07	- 16.86	0.13
C.Branco	- 26.94	0.29	5.31	0.04
Coimbra	- 11.79	0.12	28.87	0.23
Evora	3.67	0.04	3.05	0.02
Faro	2.50	0.02	2.10	0.02
Guarda	10.58	0.11	2.11	0.02
Leiria	- 93.42	0.97	0.14	0.00
Lisboa	-	-	-	-
Portalegre	- 18.39	0.20	- 5.18	0.04
Porto	2211.75	19.17	1153.20	9.16
Santarem	- 46.47	0.49	56.65	0.45
Setubal	81.35	0.81	- 5.86	0.05
V.Castelo	-107.65	1.14	8.11	0.06
V.Real	- 43.88	0.46	18.77	0.15
Viseu	-111.19	1.16	2.20	0.02

TABLE 8.27: Transition Probability Estimates for the Unrestricted OLS Using Principal Components and Stacked District Data (Method VII) - LISBOA

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.16 (4.83)	.58 (7.95)	.40 (4.40)	.33 (5.09)	.69 (8.81)	.65 (8.59)	.26 (3.10)	.89 (11.35)	-.01 (-.07)	.12 (1.33)	.93 (7.83)	.67 (5.62)	.33 (2.75)
1972	.09 (3.36)	.66 (9.51)	.33 (3.72)	.41 (6.81)	.62 (9.02)	.73 (11.11)	.19 (2.43)	.97 (14.26)	-.08 (-.94)	.20 (2.49)	.86 (7.95)	.76 (6.98)	.26 (2.43)
1973	.09 (7.32)	.68 (10.17)	.33 (4.08)	.43 (8.83)	.62 (11.15)	.75 (15.22)	.19 (3.14)	.99 (19.15)	-.07 (-1.25)	.22 (3.86)	.86 (9.53)	.77 (8.81)	.26 (2.90)
1974	.05 (1.67)	.77 (10.47)	.29 (3.22)	.52 (8.29)	.58 (7.79)	.84 (12.45)	.15 (1.81)	1.08 (13.75)	-.12 (-1.38)	.31 (3.83)	.81 (7.30)	.87 (7.83)	.22 (1.95)
1975	.16 (5.29)	.53 (7.33)	.40 (4.44)	.29 (4.65)	.68 (10.54)	.61 (9.36)	.25 (8.11)	.85 (11.59)	-.01 (-.14)	.08 (.92)	.92 (8.40)	.63 (6.96)	.33 (3.18)
1976	.12 (8.52)	.62 (9.37)	.36 (4.32)	.38 (7.53)	.64 (11.29)	.70 (13.52)	.22 (3.44)	.94 (17.65)	-.05 (-.80)	.17 (2.75)	.88 (9.53)	.72 (7.85)	.29 (3.13)
1977	.15 (14.84)	.59 (8.80)	.39 (4.83)	.34 (7.16)	.68 (12.01)	.66 (13.19)	.25 (4.26)	.90 (17.67)	-.02 (-.31)	.13 (2.32)	.91 (10.11)	.68 (7.57)	.32 (3.52)
1978	.13 (17.42)	.61 (9.16)	.37 (4.63)	.36 (7.80)	.66 (12.27)	.69 (14.54)	.23 (3.99)	.93 (18.52)	-.04 (-.64)	.16 (2.86)	.90 (10.17)	.71 (8.36)	.30 (3.45)
1979	.17 (12.60)	.56 (8.24)	.41 (4.97)	.31 (6.40)	.69 (11.82)	.64 (12.10)	.26 (4.34)	.88 (15.67)	.00 (-.03)	.11 (1.76)	.93 (9.97)	.66 (7.16)	.34 (3.61)
1980	.18 (13.01)	.53 (7.68)	.42 (5.12)	.28 (5.64)	.71 (12.48)	.60 (11.58)	.28 (4.46)	.84 (14.99)	.01 (.18)	.07 (1.15)	.94 (10.09)	.63 (7.18)	35 (3.84)
1981	.17 (8.70)	.54 (7.74)	.41 (4.79)	.29 (5.37)	.70 (10.77)	.61 (10.38)	.27 (3.86)	.86 (13.23)	.00 (.03)	.09 (1.21)	.93 (9.20)	.64 (6.25)	.34 (3.41)

Figures in parenthesis are t-values

TABLE 8.28: Transition Probability Estimates for the Unrestricted OLS Using Principal Components and Stacked District Data (Method VIII) - BRAGANCA

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.15 (.73)	.62 (2.04)	.39 (1.28)	.38 (1.25)	.68 (1.85)	.70 (1.89)	.25 (.59)	.94 (2.22)	-.01 (-.03)	.17 (.36)	.92 (1.76)	.72 (1.38)	.33 (.62)
1972	.08 (.68)	.69 (4.96)	.32 (1.69)	.44 (2.46)	.61 (2.78)	.77 (3.49)	.18 (.72)	1.01 (4.01)	-.08 (-.30)	.24 (.84)	.85 (2.67)	.79 (2.52)	.26 (.81)
1973	.13 (.94)	.64 (4.27)	.37 (1.77)	.39 (1.99)	.65 (2.72)	.71 (2.97)	.23 (.81)	.96 (3.47)	-.04 (-.13)	.19 (.60)	.89 (2.58)	.74 (2.15)	.30 (.87)
1974	.05 (.20)	.60 (2.49)	.29 (.85)	.35 (1.06)	.57 (1.42)	.67 (1.68)	.14 (.31)	.92 (1.97)	-.12 (-.24)	.15 (.28)	.81 (1.42)	.70 (1.22)	.22 (.38)
1975	.09 (.38)	.74 (2.88)	.33 (.93)	.50 (1.39)	.62 (1.43)	.82 (1.88)	.19 (.38)	1.06 (2.12)	-.07 (-.13)	.29 (.52)	.86 (1.39)	.84 (1.37)	.27 (.43)
1976	.09 (.94)	.63 (5.17)	.33 (2.05)	.38 (2.50)	.62 (3.42)	.70 (3.91)	.19 (.92)	.94 (4.54)	-.07 (-.32)	.17 (.73)	.86 (3.27)	.72 (2.78)	.27 (1.02)
1977	.17 (1.17)	.54 (3.27)	.41 (1.84)	.29 (1.33)	.70 (2.66)	.61 (2.32)	.27 (.90)	.86 (2.81)	.01 (.02)	.09 (.25)	.94 (2.52)	.64 (1.69)	.35 (.92)
1978	.17 (2.49)	.54 (5.90)	.41 (3.25)	.29 (2.91)	.70 (5.33)	.61 (4.88)	.27 (1.80)	.86 (5.96)	.01 (.04)	.09 (.69)	.94 (4.98)	.64 (3.47)	.35 (1.81)
1979	.19 (1.66)	.52 (3.93)	.43 (2.40)	.27 (1.61)	.71 (3.53)	.59 (2.93)	.28 (1.23)	.83 (3.60)	.02 (.07)	.06 (.25)	.95 (3.33)	.62 (2.12)	.36 (1.24)
1980	.18 (1.92)	.54 (4.73)	.42 (2.74)	.29 (2.09)	.70 (4.21)	.61 (3.68)	.27 (1.44)	.85 (4.49)	.01 (.04)	.08 (.40)	.94 (3.98)	.64 (2.64)	.35 (1.45)
1981	.17 (1.03)	.54 (2.92)	.42 (1.64)	.29 (1.16)	.70 (2.35)	.61 (2.02)	.27 (.79)	.85 (2.46)	.01 (.02)	.08 (.21)	.94 (2.22)	.63 (1.48)	.35 (.81)

Figures in parenthesis are t-values

TABLE 8.29: Transition Probability Estimates for the Restricted OLS Using Principal Components and Stacked District Data (Method VIII) - LISBOA

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.20	.60	.38	.26	.77	.70	.21	1.15	-.30	.06	1.01	.77	.24
	(4.36)	(6.83)	(3.46)	(3.07)	(7.58)	(7.03)	(1.91)	(10.64)	(2.50)	(.57)	(6.50)	(4.72)	(1.48)
1972	.13	.69	.30	.35	.69	.79	.13	1.24	-.38	.15	.93	.86	.17
	(3.28)	(8.28)	(2.84)	(4.58)	(8.01)	(1.30)	(1.30)	(13.27)	(3.49)	(1.41)	(6.74)	(5.83)	(1.15)
1973	.10	.73	.28	.39	.67	.83	.11	1.29	-.40	.19	.91	.90	.14
	(5.95)	(9.53)	(2.97)	(7.00)	(10.34)	(1.48)	(1.48)	(19.78)	(5.37)	(2.73)	(8.12)	(7.97)	(1.20)
1974	.03	.85	.21	.51	.60	.95	.03	1.40	-.48	.31	.84	1.02	.07
	(.74)	(9.65)	(1.92)	(6.44)	(6.42)	(10.58)	(.33)	(13.30)	(4.24)	(2.76)	(5.70)	(6.92)	(.47)
1975	.17	.57	.34	.24	.73	.68	.17	1.13	-.34	.04	.97	.74	.21
	(4.00)	(6.51)	(3.13)	(3.03)	(8.61)	(7.62)	(1.62)	(10.84)	(3.01)	(.33)	(6.88)	(5.88)	(1.43)
1976	.13	.67	.31	.34	.70	.78	.14	1.23	-.37	.14	.94	.84	.17
	(7.04)	(8.78)	(3.25)	(5.88)	(10.66)	(12.41)	(1.84)	(17.88)	(4.62)	(1.80)	(8.27)	(6.99)	(1.46)
1977	.17	.64	.34	.30	.73	.74	.17	1.20	-.34	.11	.97	.81	.21
	(11.83)	(8.44)	(3.70)	(5.71)	(11.43)	(12.81)	(12.49)	(18.61)	(4.69)	(1.48)	(8.76)	(6.91)	(1.74)
1978	.14	.67	.32	.34	.71	.77	.14	1.23	-.37	.14	.95	.84	.18
	(13.48)	(8.85)	(3.42)	(6.53)	(11.75)	(14.29)	(2.13)	(19.68)	(5.29)	(2.02)	(8.79)	(7.71)	(1.57)
1979	.17	.62	.35	.29	.74	.73	.18	1.18	-.33	.09	.98	.79	.21
	(9.41)	(8.01)	(3.73)	(5.15)	(11.03)	(11.64)	(2.46)	(16.34)	(4.29)	(1.13)	(8.51)	(6.58)	(1.77)
1980	.19	.59	.36	.25	.75	.69	.19	1.14	-.32	.05	.19	.76	.23
	(9.61)	(7.47)	(3.82)	(4.47)	(11.55)	(11.00)	(2.55)	(15.40)	(4.03)	(.64)	(8.65)	(6.57)	(1.91)
1981	.18	.61	.35	.28	.74	.71	.18	1.17	-.33	.08	.98	.78	.22
	(6.49)	(7.54)	(3.55)	(4.31)	(9.98)	(9.60)	(2.10)	(13.44)	(3.35)	(.80)	(7.86)	(5.65)	(1.71)

Figures in parenthesis are t-values

TABLE 8.30: Transition Probability Estimates for teh Restricted OLS Using Principal Components and Stacked District Data (Method VIII) - BRAGANCA

YEAR	P_{55}	P_{56}	P_{66}	P_{67}	P_{77}	P_{78}	P_{88}	P_{89}	P_{99}	$P_{9,10}$	$P_{10,10}$	$P_{10,11}$	$P_{11,11}$
1971	.15 (.53)	.66 (2.14)	.33 (.79)	.33 (.76)	.72 (1.43)	.76 (1.46)	.16 (.27)	1.22 (2.02)	-.35 (.54)	.13 (.19)	.96 (1.34)	.83 (1.12)	.19 (.27)
1972	.09 (.53)	.73 (3.80)	.27 (1.04)	.40 (1.55)	.66 (2.20)	.84 (2.68)	.09 (.27)	1.29 (3.58)	-.42 (1.08)	.20 (.49)	.90 (2.08)	.90 (2.01)	.13 (.30)
1973	.13 (.71)	.68 (3.30)	.31 (1.11)	.35 (1.21)	.70 (2.13)	.79 (2.31)	.14 (.36)	1.24 (3.17)	-.37 (.88)	.15 (.34)	.94 (1.99)	.85 (1.75)	.17 (.37)
1974	.05 (.16)	.66 (1.94)	.23 (.50)	.33 (.70)	.62 (1.11)	.77 (1.33)	.06 (.09)	1.22 (1.83)	-.45 (.63)	.13 (.17)	.86 (1.08)	.83 (1.02)	.09 (.12)
1975	.11 (.31)	.78 (2.12)	.29 (.57)	.45 (.87)	.68 (1.11)	.89 (1.41)	.11 (.16)	1.34 (1.85)	-.40 (.51)	.25 (.31)	.92 (1.06)	.95 (1.07)	.15 (.17)
1976	.11 (.77)	.68 (4.17)	.29 (1.31)	.35 (1.65)	.68 (2.72)	.79 (3.09)	.11 (.39)	1.24 (4.20)	-.40 (1.24)	.15 (.45)	.92 (2.54)	.85 (2.31)	.15 (.41)
1977	.19 (.90)	.62 (2.57)	.36 (1.19)	.26 (.83)	.75 (2.08)	.70 (1.83)	.19 (.45)	1.15 (2.62)	-.32 (.60)	.06 (.12)	.99 (1.93)	.77 (1.41)	.23 (.43)
1978	.19 (1.96)	.62 (4.97)	.37 (2.24)	.29 (1.89)	.76 (4.27)	.72 (3.95)	.19 (.95)	1.18 (5.63)	-.32 (1.42)	.09 (.38)	1.00 (3.90)	.79 (2.94)	.23 (.88)
1979	.20 (1.30)	.58 (3.19)	.38 (1.58)	.24 (1.02)	.77 (2.78)	.68 (2.35)	.21 (.65)	1.13 (3.41)	-.13 (.87)	.04 (.12)	.101 (2.57)	.75 (1.80)	.24 (.61)
1980	.19 (1.53)	.60 (3.93)	.37 (1.85)	.26 (1.36)	.76 (3.37)	.70 (2.99)	.20 (.76)	1.15 (4.28)	-.31 (1.10)	.06 (.21)	1.00 (3311)	.77 (2.26)	.23 (.71)
1981	.20 (.84)	.60 (2.33)	.37 (1.09)	.27 (.78)	.77 (1.87)	.71 (1.63)	.20 (.43)	1.16 (2.32)	-.31 (.59)	.07 (.12)	1.00 (1.73)	.77 (1.25)	.24 (1.41)

Figures in parenthesis are t-values

show that the transition probability estimates corresponding to terminal and first grades of levels of education ($\hat{p}_{67}$, $\hat{p}_{77}$, $\hat{p}_{9,10}$ and $\hat{p}_{10,10}$) are very low for promotion probabilities and very high for repetition probabilities.

Although the t-values for the transition probability estimates are better for Lisboa than for Braganca, the unrestricted OLS estimator has still produced non-significant t-values at the 5% level for $\hat{p}_{99}$ and the restricted OLS estimator has produced non-significant t-values at the same level for the estimates $\hat{p}_{88}$, $\hat{p}_{9,10}$ and $\hat{p}_{11,11}$.

The results of performing the regression of the observed point estimates on these transition probability estimates, that is, when stacked district data and principal components are used, are presented in Table 8.31 for the two districts analysed. The table shows that the restricted OLS estimation procedure has generated estimates better correlated with the observed point estimates for both districts, the district of Lisboa, however, giving serial correlatión for two of the regressions performed.

The previous sections have shown the unrestricted stepwise OLS estimation procedure applied simultaneously to all stacked district data (Method III) to be the estimation procedure which has produced more reasonable transition probability estimates and similar time-patterns for these estimates. It seems worthwhile, then, to compare these results with those obtained when principal components were used. Similar to what was observed for the whole country, the use of principal components as the explanatory variables of the model has not improved the results; Method III (unrestricted OLS) still gives more reliable transition probability estimates, and with

TABLE 8.31: Comparison between the Estimates Obtained Using Principal Components and Stacked District Data (Method VIII)

	LISBOA								BRAGANCA							
	Unrestricted OLS				Restricted OLS				Unrestricted OLS				Restricted OLS			
	Mean Value	Stand Dev.	R[(a)]	D-W	Mean Value	Stand Dev.	R[(a)]	D-W	Mean Value	Stand Dev.	R	D-W	Mean Value	Stand Dev.	R	D-W
P_{55}	.1337	.0410	.75649	1.80	.1456	.0481	.78488	1.80	.1351	.0484	.74038	1.25	.1468	.0515	.73858	1.14
P_{56}	.6055	.0726	.64856	1.42	.6567	.0783	.63311	1.39	.6019	.0724	.64872	1.57	.6532	.0639	.67346	1.65
P_{66}	.3741	.0408	.66020	2.55	.3235	.0481	.73248	2.73	.3752	.0484	.59336	2.10	.3246	.0515	.57633	1.99
P_{67}	.3570	.0726	.63701	1.70	.3239	.0783	.57392	1.45	.3530	.0724	.73047	1.36	.3195	.0639	.75824	1.60
P_{77}	.6605	.0408	.38465	.67*	.7145	.0481	.39358	.73*	.6617	.0484	.65766	1.94	.7154	.0515	.63278	1.91
P_{78}	.6795	.0726	.46763	1.85	.7607	.0783	.48600	1.81	.6757	.0724	.50793	2.02	.7572	.0639	.51057	1.92
P_{88}	.2315	.0408	.59058	.75*	.1494	.0481	.64164	.87*	.2326	.0484	.79215	2.30	.1504	.0515	.79065	2.20
P_{89}	.9215	.0726	.55440	1.46	1.2149	.0783	.63810	1.50	.9179	.0724	.26391	2.08	1.2115	.0639	.32863	2.19
P_{99}	-.0345	.0408	.53696	2.22	-.3615	.0481	.58282	2.18	-.0334	.0484	.63275	2.11	-.3605	.0515	.65714	2.13
$P_{9,10}$	.1515	.0726	.58645	1.45	.1234	.0783	.60371	1.43	.1479	.0724	.32365	1.48	.1199	.0639	.40001	1.58
$P_{10,10}$	.8976	.0408	.18877	1.97	.9526	.0481	.32426	1.92	.8988	.0484	.61731	2.67	.9539	.0515	.64479	2.83
$P_{10,11}$	.7042	.0726	.22423	2.25	.8284	.0783	-.18710	2.20	.2005	.0724	-.18826	1.08	.8249	.0639	-.12720	.99*
$P_{11,11}$	.3055	.0408	.54879	1.42	.1855	.0481	.45340	1.15	.3066	.0724	.2911	2.19	.1864	.0515	.31326	2.20

(a) R is the correlation coefficient between the observed point estimates and the transition probability estimates.

time-patterns more similar to the time-patterns of the observed point estimates.

8.6 Final Remarks

The regional level application of the extended Markov model to the Portuguese educational system, has shown the unrestricted stepwise OLS estimation procedure, using all stacked district data (Method III), to be the estimation method that has globally produced more reasonable results. However, the impossibility of performing the GLS estimation with the stacked district data has made it impossible to assess the gains that would be obtained from this procedure. Nevertheless, some remarks can be made on the pooled cross-section time-series estimation procedure performed by regions. The process has led to an improvement in the reliability of the estimates of the transition probabilities for Region 2, as low standard errors and significant t-values at the 5% level have occured. Also, when the regression analyses were performed between the observed point estimates and the corresponding transition probability estimates, the GLS estimator has proved to give better correlation coefficients for Region 2. The same, however, cannot be asserted for Region 1: although the t-values are slightly better when the pooled cross-section time-series estimation procedure is performed, the correlation coefficients and, therefore, the time-patterns are not as good. This suggests that if the unrestricted GLS estimator could be applied to the stacked district data, more reliable estimates would occur but not much better patterns, in general, would result. Unreasonable transition probability estimates associated with the terminal or first grades of the education levels would also result.

Chapter 8

Footnotes

1. Mutually independent disturbance terms is a questionable assumption when the cross-sectional units are geographical regions. KMENTA (p.512) suggests the cross-sectionally correlated and time-series autoregressive model. This attempt has not been made because the results of the pooled cross-sectionally heteroscedastic and time-series autoregressive model has been shown to be not much better than the previous methods, especially for Region 1.

2. As well as for the case of the unrestricted OLS estimator, the repetition probability estimates for each grade are equal for all districts of Region 1.

3. Note that there is no significant difference between the unrestricted OLS and the unrestricted GLS (pooled cross-section time-series estimation) estimates of the coefficients of the model.

4. Note that the factor loading for variable PUCLASS for Portalegre is also positive.

Chapter 9

ANALYSIS OF THE RESULTS AND POLICY IMPLICATIONS

9.1. Introduction

The application of the basic Markov model to the preparatory and secondary levels of education of the Portuguese educational system, developed in Chapters 4 and 5, has shown, as expected, the weakness of the assumption of stationary transition probabilities, thus indicating the gap between the simple model and the school enrolment process. This result is consistent with the diverse comments made by specialists in educational planning and previously referred to in the literature review in Chapter 2.

The assumption of stationary transition probabilities was, therefore, subsequently replaced by a more flexible assumption, allowing the changes in the transition probabilities to be affected by changes in explanatory variables. An extended Markov model was built incorporating behavioural relationships for the promotion and the repetition probabilities. The application of the extended Markov model to the Portuguese educational system was performed in Chapters 7 and 8, the transition probabilities being estimated not only for the whole country but also by district. Different attempts were made to find the estimates of the transition probabilities. The time-varying transition probability estimates obtained show time-patterns very much like the time-patterns of the corresponding observed point estimates, for most of the methods used, and for most of the cases. The results obtained give reassurance of the validity of incorporating a causal structure into the traditional Markov model. Of all the estimation procedures undertaken, the unrestricted stepwise OLS estimation method applied simultaneously to all stacked district data (Method III) is the one which has globally produced most reasonable transition probability estimates; that is, least absolute differences, best t-values, no serial correlations of the residuals, and transition probability estimate time-patterns most similar to the corresponding observed point estimate patterns.

The explanatory variables used to analyse the changes in the repetition and promotion probabilities were grouped into supply side and demand side factors and are described in Table 7.2. As described in Chapter 8, only nine out of the seventeen explanatory variables selected display different values by district (PUTEA, PUCLASS, BUS, HELP, UNQUAL, LIFE, ILLIT, UNEMP, EARN) and can be considered potentially to explain some of the regional disparities observed in the behaviour over time of the transition probabilities. Among all the factors used to explain the changes in the probabilities, there are two that appear as being significant in the changes of the promotion probabilities for all the estimation methods performed. These are the percentage of students who get a scholarship from the social services (HELP), and the number of unemployed workers (UNEMP).

9.2. The Behavioural Relationships and their Implications

9.2.1. The Whole Country

The behavioural equations resulting from the application of Method III (unrestricted stepwise OLS estimation using all stacked district data) to the Portuguese educational system are as follows:

$$p_{ii} = a_{ii} + \underset{(2.98)}{0.0154}\ \text{EDUC} - \underset{(3.94)}{0.0250}\ \text{PUCLASS} + \underset{(5.62)}{0.0354}\ \text{GDP} + v_{ii} \quad (9.1)$$

$$\begin{aligned} p_{i,i+1} = a_{i,i+1} & - \underset{(2.75)}{0.0269}\ \text{PCAP} - \underset{(6.20)}{0.0326}\ \text{TEARN} - \underset{(2.12)}{0.0273}\ \text{PUTEA} \\ & + \underset{(4.34)}{0.0352}\ \text{HELP} + \underset{(3.30)}{0.0597}\ \text{GDP} - \underset{(7.94)}{0.0726}\ \text{UNEMP} \\ & + \underset{(4.93)}{0.0387}\ \text{EARN} - \underset{(5.44)}{0.1070}\ \text{LFLEV} + v_{i,i+1} \quad (9.2) \end{aligned}$$

(t-values in parentheses)

As the explanatory variables have been included in the model in a standardised form (zero mean, unit variance), the values of the a_{ii} and $a_{i,i+1}$ terms can be identified as the mean values of the transition probability estimates. In other words, the constant terms represent the estimates of the transition probabilities when the explanatory variables tahe their mean values; that is, if the educational system could be described by the basic Markov model.

The regression coefficients specify the individual effect of each explanatory variable upon the transition probabilities while holding the other explanatory variables constant. Equation (9.1) shows that only variable PUCLASS resulted as a determinant on the district differences of the repetition probabilities. Equation (9.2) reveals that the promotion probabilities are more sensitive to external factors than the repetition probabilities, the unemployment and the level of education of the labour force being the factors that individually show more effect on the changes of the promotion probabilities. For these probabilities, two supply side explanatory variables (PUTEA, HELP) and two demand side explanatory variables (UNEMP, EARN) appeared to be the factors that explain the district disparities.

As the behavioural relationships introduced in the extended model are linear, the estimated coefficient of each explanatory variable represents its marginal effect in the corresponding transition probabilities. Table 9.1 reports for each behavioural relationship estimated and for each significant explanatory variable its marginal effect in the corresponding transition probabilities.

Table 9.1 Determinants of the Transition Probability Changes And Their Marginal Effect - Whole Country

Transition Probability	Explanatory Variable	Coeff. of Stand. Variable	Coeff. of Unstand. Variable
p_{ii}	EDUC	0.0154	0.0003
	PUCLASS	-0.0250	-0.0065
	GDP	0.0354	0.0004
$p_{i,i+1}$	PCAP	-0.0269	-0.0110
	TEARN	-0.0326	-0.0012
	PUTEA	-0.0273	-0.0208
	HELP	0.0352	0.0050
	GDP	0.0597	0.0029
	UNEMP(x1000)	-0.0726	-0.0352
	EARN(x100)	0.0387	0.0280
	LFLEV	-0.1070	-0.0416

For a better understanding of the magnitude of the effect produced, the estimated coefficients of each unstandardised explanatory variable are also presented in the table. These values have been derived by applying a linear transformation to the equations that relate the transition probabilities with the explanatory variables described as follow

$$\hat{p}^t_{ii} = \hat{a}_{ii} + \underline{x}^{t'} \hat{\underline{\delta}}_r + v^t_{ii}$$

$$\hat{p}^t_{i,i+1} = \hat{\underline{a}}_{i,i+1} + \underline{x}^{t'} \hat{\underline{\delta}}_p + v^t_{i,i+1}$$

The disturbances originated by using the original explanatory variables or by using their standardised form are the same. Identifying the standardised explanatory variables by $x^u_h(t)$ and the unstandardised explanatory variables by $x^s_h(t)$, the following identities can be obtained

$$a^u_{ii} + \underline{x}^{t'u} \hat{\underline{\delta}}^u_r = a^s_{ii} + \underline{x}^{t's} \hat{\underline{\delta}}^s_r$$

$$a^u_{i,i+1} + \underline{x}^{t'u} \hat{\underline{\delta}}^u_p = a^s_{i,i+1} + \underline{x}^{t's} \hat{\underline{\delta}}^s_p$$

Thus, the $\underline{a}$ and $\underline{\delta}$ parameters corresponding to the unstandardised explanatory variables can be estimated by solving a system of simultaneous equations. It is apparent from the table that the strongest marginal effects are produced by variable PUCLASS upon the repetition probabilities and by variable LFLEV upon the promotion probabilities. However, when trying to compare the responsiveness of the transition probabilities to changes in the explanatory variables, these variables have different units of measurement. These problems of dimension can be overcome by using the corresponding elasticities at the mean. The elasticity at the mean of the transition probability is defined as $(\Delta p/p)/(\Delta\bar{x}/\bar{x})$ where $\Delta p/p$ is the proportionate change in the transition probability and $\Delta\bar{x}/\bar{x}$ is the proportionate change in the mean of the explanatory variable. The results obtained in the previous chapter show that the application of the extended Markov model to the Portuguese educational system has produced transition probability estimates with patterns most like the patterns of the observed point estimates. However, some of the estimates present, in absolute terms, large differences when compared with the corresponding observed point estimates. Thus, it seems worthwhile for further policy to present the elasticities of the observed point estimates at the mean of each explanatory variable, instead of calculating the elasticities of the transition probability estimates, because those are the values that reflect the movements of the students within the educational system for the time period of analysis. The values of the different elasticities are presented in Tables 9.2 and 9.3, which also show the percentage of the transition probability estimate at the mean that is explained either by the constant term (A) or by the set of explanatory variables (B), using the usual result that the estimators exactly satisfy the estimated equation at the mean of the variables. Table 9.2 shows that the higher elasticity of the repetition probabilities are given by the variable PUCLASS. Table 9.3 also shows that variable UNEMP is the one that has the highest elasticity of the promotion probabilities. A one per cent decrease in the average number of unemployed workers will lead to an average increase of nine per cent in the promotion probability, so there is a very high responsiveness to changes in variable UNEMP.

However, the characteristics of the demand side factors, trying to explain the socio-economic influences in the students performance and students attitudes towards pursuing their studies, show that their effect in producing a change in the transition probabilities must be seen to be having a long-term effect. Thus, although the demand side variables elasticities of the transition probabilities are higher than the supply side variables elasticities of the transition probabilities, its implementation cannot be easy. Any

Table 9.2 Percentage of The Repetition Probability Estimate Explained by The Explanatory Variables and Elasticities at The Mean -WHOLE COUNTRY

Prob.	A(%)	B(%)	Elasticity EDUC	Elasticity PUCLASS	Elasticity GDP
P_{55}	237.24	-17.24	0.03	-2.06	0.58
P_{66}	173.19	-73.19	0.03	-2.06	0.58
P_{77}	131.37	-31.37	0.02	-1.55	0.43
P_{88}	161.54	-61.54	0.02	-1.47	0.41
P_{99}	269.46	-169.46	0.02	-1.24	0.34
$P_{10,10}$	122.98	- 22.98	0.06	-4.43	1.23
$P_{11,11}$	136.60	- 36.60	0.02	-1.63	0.45

Table 9.3 Percentage of The Promotion Probability Estimate Explained by The Explanatory Variables and Elasticities at The Mean -WHOLE COUNTRY

Prob.	A(%)	B(%)	Elasticity PCAP	TEARN	PUTEA	HELP	GDP	UNEMP	EARN	LFLEV
P_{56}	125.72	-25.72	-0.22	-0.01	-0.42	0.10	0.86	-8.95	0.73	-0.94
P_{67}	157.96	-57.96	-0.23	-0.10	-0.45	0.10	0.91	-9.47	0.77	-0.99
P_{78}	133.68	-33.68	-0.23	-0.01	-0.45	0.10	0.91	-9.47	0.77	-0.99
P_{89}	125.56	-25.56	-0.23	-0.01	-0.45	0.10	0.91	-9.47	0.77	-0.99
$P_{9,10}$	148.48	-48.48	-0.25	-0.01	-0.49	0.11	1.09	-10.36	0.84	-0.99
$P_{10,11}$	94.93	- 5.07	-0.19	-0.01	-0.38	0.09	0.76	-7.97	0.65	-0.84

A = Percentage of the transition probability estimate explained by the constant term a_{ij} at the mean of the observed point estimates.

B = Percentage of the transition probability estimate explained by the set of the explanatory variabes at the mean of the observed point estimates.

change in a supply side explanatory variable is easier to effect as it depends exclusively on the policy measures proposed by the Ministry of Education, and on the availability of resources to implement such measures.

A closer examination of the signs of the coefficients of the selected explanatory variables must be made to clarify the meaning of the behavioural relationships. However, the existence of multicollinearity between the set of explanatory variables must be recalled at this stage. The major undesirable consequence of multicollinearity is that the variances of the OLS estimates of the parameters of correlated variables are quite large. It is the uncertainty as to which explanatory variable deserves the credit for the jointle explained variation in the dependent variable that creates the uncertainty as to the true values of the coefficients being estimated and thus causes the higher variances of their estimates. The parameter estimates are, therefore, not precise. This suggests that although some conclusions are presented in the study, they must be regarded carefully and their validation should be further explored.

The signs of the coefficients of the supply side explanatory variables PUTEA and HELP reveal that the smaller the pupil-teacher ratio, the better achievement is observed and an increase in the promotion probability results. More precisely, a decrease of one per cent in the average pupil-teacher ratio will lead to almost a one half per cent increase in the promotion probability for most of the grades. Also, an increase of one per cent in the facilities offered to the students, such as scholarships, books and lodging allowances, results in an increase of one tenth of a per cent in the promotion probability. The noticeable strong correlation between variable PUTEA and variable UNQUAL must be noted at this stage, implying that a decrease in the percentage of teachers without qualification may have a positive effect in the changes of the promotion probabilities.

One rather surprising finding is that BUS is not a significant variable regarding the repetition or the promotion probabilities. Although this variable has been eliminated during the regression, it is strongly negatively correlated with the pupil-teacher ratio. Therefore, one can infer that BUS also may effect the changes in the promotion probabilities. An increase in the bussing facilities allowing the students to attend the school, may produce an increase in the promotion probabilities.

The strong influence of the unemployment rate and the average earnings level on the students' decision on pursuing their studies has been pointed out in the literature [see UNESCO, 1979; p.83]. Once more, the behavioural relationships estimated confirm the statement, showing a negative influence of the number of unemployed workers in the changes of the promotion probabilities. Also, variable EARN appears with a positive coefficient, implying that an increase in the level of the average earnings produces an increase in the promotion probabilities. By identifying higher salaries with higher levels of qualification, one can infer that an increase in variable EARN means an increase in the educational level of the population, together indicating the students' interest in pursuing their studies.

The percentage of capital expenditures in the education expenditures (PCAP) can be a measure of deficient conditions offered to the students. High values observed for variable PCAP may be associated with overcrowded schools, or lack of equipment such as, for instance, the necessity of building new schools or buying school equipment. The negative sign of the parameter estimate associated with this explanatory variable reveals, therefore, that a one per cent decrease in the average percentage of capital expenditures implies an average increase of one quarter of a per cent in the promotion probabilities. In this context, a decrease in the percentage of capital expenditures should be interpreted as an indication of the improved physical resources offered to the students.

An examination of equation (9.1) and equation (9.2) show that some unexpected results were obtained. Whereas, for example, a positive sign may be expected for variable PUCLASS, the parameter estimate is negative. However, it must be noted that the use of the same school room by more than one group of students in several schools, and the non-accounting of this fact by this explanatory variable, has given a distorted meaning to this variable. A high pupil-classroom ratio can denote double-use and may correspond, therefore, to a low ratio in reality.

Also unexpected is the negative sign of variable TEARN, implying that an increase in the average teachers' earnings would be reflected in a decrease in the promotion probabilities. As this variable is not correlated with any other explanatory variable, no straightforward explanation about the negative sign can be found. The strong changes in the structure and pay scale of the teachers' career that took place in the middle of the 1970s may, however, be the origin of this result.

Finally, the positive signs of variables EDUC and GDP reveal that the cause-effect relationship is not very helpful for these variables. For example, the positive sign of the coefficient of variable EDUC indicates direct variation between the repetition probabilities and the percentage increase in the expenditures on education. But how do the repetition probabilities change? It seems, however, to be the increase in the repetition probabilities that increases the number of students in the system and so requires more expenditure on education. The cause-effect relationship should then be the other way about, suggesting, therefore, that in further studies this variable can be removed from the set of explanatory variables. In contrast to variable EDUC, variable GDP is highly correlated with other demand side factors. In particular, GDP is highly positively correlated with the percentage of the labour force whose educational attainment is either preparatory or secondary levels (LFLEV). This variable appears in the behavioural relationship for the promotion probabilities [equation 9.2] with a negative sign, whereas GDP appears in the same equation with a positive sign. This situation suggests that the influence of the changes of one of these explanatory variables on the changes in the promotion probabilities must be weighted using the influence of the other variable.

9.2.2. The Regions and Districts

In order to improve the knowledge of the effect of the explanatory variables in the changes of the transition probabilities, the extended Markov model was applied at the regional level. Of the estimation procedures applied, the unrestricted OLS stepwise regression method, performed after selecting the significant supply side and demand side explanatory variables (Method V), has proved to be the one that has produced most reasonable estimates for the transition probabilities. As described in Chapter 8, two regions have been studied, Region 1 comprising the more industrialised districts and Region 2 with mainly rural characteristics. The results, however, have shown that the application of the extended Markov model to Region 2 has produced more reasonable estimates than the results obtained for Region 1. These results suggest that the characteristics of Region 2 are more homogeneous than when moving to the more industrialized districts. In these districts the differences between urban and rural areas are more emphasized.

The behavioural equations estimated using Method V are as follows:

Region 1

$$p_{ii} = a_{ii} \underset{(3.27)}{- 0.031\ \mathrm{ILLIT}} \underset{(2.11)}{- 0.023\ \mathrm{EARN}} + u_{ii} \qquad (9.3)$$

$$\begin{aligned} p_{i,i+1} = a_{i,i+1} & \underset{(2.83)}{- 0.024\ \mathrm{EDUC}} \underset{(4.67)}{- 0.040\ \mathrm{TEARN}} \underset{(3.30)}{+ 0.035\ \mathrm{HELP}} \\ & \underset{(2.33)}{+ 0.042\ \mathrm{UNQUAL}} \underset{(2.59)}{- 0.036\ \mathrm{UNEMP}} \\ & \underset{(3.07)}{+ 0.036\ \mathrm{EARN}} + u_{i,i+1} \qquad (9.4) \end{aligned}$$

Region 2

$$\begin{aligned} p_{ii} = a_{ii} & \underset{(2.68)}{+ 0.018\ \mathrm{TEARN}} \underset{(2.51)}{+ 0.028\ \mathrm{HELP}} \underset{(2.63)}{- 0.031\ \mathrm{UNQUAL}} \\ & \underset{(1.97)}{- 0.019\ \mathrm{EARN}} + u_{ii} \qquad (9.5) \end{aligned}$$

$$
\begin{aligned}
p_{i,i+1} = a_{i,i+1} & + \underset{(4.40)}{0.025}\ \mathrm{EDUC} - \underset{(9.01)}{0.069}\ \mathrm{TEARN} - \underset{(4.57)}{0.045}\ \mathrm{PUTEA} \\
& + \underset{(5.64)}{0.066}\ \mathrm{HELP} + \underset{(6.16)}{0.095}\ \mathrm{UNQUAL} - \underset{(3.50)}{0.031}\ \mathrm{UNEMP} \\
& + \underset{(2.79)}{0.027}\ \mathrm{EARN} + u_{i,i+1} \qquad (9.6)
\end{aligned}
$$

(t-values in parentheses)

It seems to be true for the several methods performed that the promotion probabilities are more sensitive to external factors than the repetition probabilities. Equation (9.3) shows that for the industrialized region, only demand side explanatory variables resulted as determinants in the changes of the repetition probabilities. In contrast, equation (9.5) reveals that the repetition probabilities for the region with rural characteristics is more sensitive to the changes in the supply side factors. The behavioural equations estimated for the promotion probabilities are very similar, the difference being given by variable PUTEA which appears significant for Region 2. Although the available data do not show the internal disparities within each region (between urban and rural areas), one main general finding results from the analyses: Region 2 is seen to be more sensitive to external factors than Region 1.

An interesting finding is the positive influence of variable HELP in the changes of both repetition and promotion probabilities for Region 2. One can infer from this result that an increase in the scholarships, books and lodging allowances produces a decrease in the overall drop-out probabilities for the region with more rural characteristics.

The comparison between the different behavioural equations estimated confirm the statement that the relationship between the education system and the regional socio-economic environment is complex, this relationship being subject to many factors such as the respective evolution of educational structures and economic structures, or the configuration of the educational network, itself linked to cultural traditions or to past and present regional employment prospects. Therefore, at a regional level, the functions that education performs within the national context are disturbed by migration and exchanges, and characterised by their own economic structures with varying rates of change and unequal development levels.

The marginal effect on the repetition and promotion probabilities corresponding to both the standardised or the unstandardised explanatory variables are presented in Tables 9.4 and 9.5. Tables 9.6 to 9.9 display the elasticities at the mean of the different observed point estimates for both regions. A comparison between the tables reveals that the responsiveness of the transition probabilities to changes in the level of earnings of the population is much higher for Region 1 than for Region 2, for both repetition and promotion probabilities. Also, a change in the number of unemployed workers leads to a very strong response by the promotion probabilities for the more industrialized region. In contrast, it is the percentage of unqualified teachers that displays the highest elasticities for the region with more rural characteristics.

It must also be noted that although the restricted OLS stepwise estimation procedure applied to the aggregate data for the whole country (Method I), developed in Chapter 7, has produced no reliable

Table 9.4 Determinants of the Transition Probability Changes And Their Marginal Effect REGION 1

Transition Probability	Explanatory Variable	Coeff. of Stand. Variable	Coeff. of Unstand. Variable
p_{ii}	ILLIT	-0.031	-0.0004
	EARN(x100)	-0.023	-0.0046
$p_{i,i+1}$	EDUC	0.024	0.0003
	TEARN	-0.040	-0.0022
	HELP	0.037	0.0073
	UNQUAL	0.042	0.0025
	UNEMP(x1000)	-0.036	-0.0344
	EARN(x100)	0.036	0.0098

Table 9.5 Determinants of the Transition Probability Changes And Their Marginal Effect REGION 2

Transition Probability	Explanatory Variable	Coeff. of Stand. Variable	Coeff. of Unstand. Variable
p_{ii}	TEARN	0.018	0.0018
	HELP	-0.028	-0.0013
	UNQUAL	-0.031	-0.0003
	EARN(x100)	-0.019	-0.0068
$p_{i,i+1}$	EDUC	0.025	0.0023
	TEARN	-0.069	-0.0060
	PUTEA	-0.045	-0.0147
	HELP	0.066	0.0069
	UNQUAL	0.095	0.1107
	UNEMP(x1000)	-0.031	-0.0443
	EARN(x100)	0.027	0.0007

Table 9.6 Percentage of The Repetition Probability Estimate Explained by The Explanatory Variables and Elasticities at The Mean -REGION 1

Prob.	A(%)	B(%)	Elasticity ILLIT	Elasticity EARN
P_{55}	124.66	-24.66	-0.06	-6.96
P_{66}	123.43	-23.43	-0.06	-7.42
P_{77}	110.45	-10.45	-0.04	-5.57
P_{88}	112.29	-12.29	-0.04	-5.30
P_{99}	123.05	-23.05	-0.04	-4.64
$P_{10,10}$	104.82	- 4.82	-0.01	-1.59
$P_{11,11}$	116.49	-16.49	-0.05	-5.86

Table 9.7 Percentage of The Promotion Probability Estimate Explained by The Explanatory Variables and Elasticities at The Mean -REGION 1

Prob.	A(%)	B(%)	Elasticity EDUC	TEARN	HELP	UNQUAL	UNEMP	EARN
P_{56}	76.28	23.72	0.01	-0.01	0.14	0.08	-70.15	2.25
P_{67}	66.26	33.74	0.01	-0.01	0.14	0.08	-73.11	2.24
P_{78}	79.76	20.24	0.01	-0.01	0.15	0.08	-77.48	2.49
P_{89}	82.24	17.76	0.01	-0.01	0.15	0.08	-75.23	2.42
$P_{9,10}$	71.64	28.36	0.01	-0.01	0.16	0.09	-81.11	2.61
$P_{10,11}$	48.91	51.09	0.01	-0.01	0.12	0.07	-64.88	2.09

A = Percentage of the transition probability estimate explained by the constant term a_{ij} at the mean of the observed point estimates.

B = Percentage of the transition probability estimate explained by the set of the explanatory variabes at the mean of the observed point estimates.

Table 9.8 Percentage of The Repetition Probability Estimate Explained by The Explanatory Variables and Elasticities at The Mean -REGION 2

Prob.	A(%)	B(%)	Elasticity TEARN	HELP	UNQUAL	EARN
P_{55}	113.99	-13.99	0.05	-0.18	-0.08	-1.10
P_{66}	106.96	- 6.96	0.04	-0.17	-0.07	-1.03
P_{77}	108.33	- 8.33	0.03	-0.13	-0.05	-0.77
P_{88}	109.66	- 9.66	0.03	-0.11	-0.05	-0.67
P_{99}	113.16	-13.16	0.02	-0.09	-0.04	-0.55
$P_{10,10}$	104.64	- 4.64	0.10	-0.42	-0.18	-2.57
$P_{11,11}$	119.79	-19.79	0.03	-0.12	-0.05	-0.74

Table 9.9 Percentage of The Promotion Probability Estimate Explained by The Explanatory Variables and Elasticities at The Mean -REGION 2

Prob.	A(%)	B(%)	Elasticity EDUC	TEARN	PUTEA	HELP	UNQUAL	UNEMP	EARN
P_{56}	120.01	-20.01	0.01	-0.03	0.29	0.18	5.38	-2.11	0.02
P_{67}	126.64	-26.64	0.05	-0.04	-0.37	0.23	6.84	-2.68	0.03
P_{78}	110.64	-10.64	0.05	-0.03	-0.33	0.20	6.02	-2.36	0.02
P_{89}	103.32	- 3.32	0.04	-0.03	-0.31	0.19	5.68	-2.23	0.02
$P_{9,10}$	115.12	-15.12	0.05	-0.04	-0.38	0.23	6.95	-2.72	0.03
$P_{10,11}$	103.73	- 3.73	0.04	-0.03	-0.27	0.16	4.98	-1.95	0.02

A = Percentage of the transition probability estimate explained by the constant term a_{ij} at the mean of the observed point estimates.

B = Percentage of the transition probability estimate explained by the set of the explanatory variabes at the mean of the observed point estimates.

TABLE 9.10: Estimated Values of the a Coefficients for the Restricted OLS Estimation Procedure (Method I)

COEFF	AVEIRO	BEJA	BRAGA	BRAGANCA	C.BRANCO	COIMBRA	EVORA	FARO	GUARDA
a_{55}	.142 (.0117)	.140 (.0157)	.155 (.0110)	.113 (.0177)	.139 (.0143)	.139 (.0145)	.167 (.0218)	.164 (.0190)	.133 (.0139)
a_{56}	.473 (.2424)	.806 (3.084)	.323 (.1905)	.716 (.2264)	.378 (.2045)	.677 (.2549)	.506 (.2473)	-.026 (.2651	.737 (.1589)
a_{66}	.506 (.2152)	.059 (.3707)	.619 (.2720)	.170 (.0332)	.602 (.2390)	.256 (.3042)	.438 (.2868)	1.082 (.2986)	.171 (.1954)
a_{67}	.603 (.1235)	.871 (.2523)	.460 (.1211)	.987 (.1800)	.367 (.2162)	.780 (.1988)	.370 (.2294)	.905 (.2108)	.674 (.1932)
a_{77}	.156 (.1849)	-.195 (.3569)	.327 (.1933)	-.186 (.2259)	.513 (.3125)	.113 (.2370)	.577 (.2986)	.025 (.2391)	-.008 (.3013)
a_{78}	.626 (.1015)	.542 (.1184)	.494 (.1078)	.610 (.1069)	.502 (.1516)	.513 (.1229)	.602 (.2042)	.311 (.1313)	.672 (.1604)
a_{88}	.262 (.1251)	.336 (.1533)	.410 (.1377)	.291 (.1305)	.468 (.1689)	.407 (.1482)	.311 (.2453)	.657 (.1577)	.295 (.1786)
a_{89}	.909 (.1391)	.897 (.1469)	.748 (.1393)	.997 (.1461)	.503 (.1156)	.836 (.1396)	.712 (.2252)	.607 (.1916)	1.089 (.1668)
a_{99}	-.075 (.1712)	.017 (.1679)	.121 (.1739)	.082 (.1420)	.490 (.1292)	.106 (.1553)	.329 (.2320)	.379 (.2138)	-.018 (.1644)
$a_{9,10}$	.376 (.1511)	.255 (.1173)	.132 (.2117)	.215 (.1426)	.529 (.1625)	.202 (.1325)	.708 (.1884)	.498 (.1777)	.427 (.1135)
$a_{10,10}$	.421 (.2696)	.482 (.2786)	.888 (.3477)	.582 (.3112)	.186 (.2790)	.768 (.2214)	-.246 (.3573)	.251 (.2938)	.404 (.1836)
$a_{10,11}$	.634 (.2416)	.838 (.2330)	.872 (.2950)	.416 (.3399)	.317 (.2403)	.426 (.2246)	.269 (.2598)	.807 (.1507)	.756 (.1617)
$a_{11,11}$	.473 (.2424)	.266 (.2426)	.194 (.3201)	.717 (.3822)	.811 (.263)	.659 (.2413)	.858 (.2769)	.193 (.1673)	.364 (.1631)
R^2	.9989	.9985	.9990	.9986	.9990	.9992	.9978	.9990	.9992
$\bar{R}^2$	.9985	.9977	.9984	.9979	.9985	.9988	.9968	.9985	.9987
F value	2013.22	1182.03	1783.69	1410.41	1820.52	2299.93	893.53	1920.65	2243.14
df	58	55	56	58	55	56	57	56	57

Figures in parenthesis are standard errors

TABLE 9.10: Estimated Values of the a Coefficients for the Restricted OLS Estimation Procedure (Method I) (Continued)

COEFF	LEIRIA	LISBOA	PORTALEGRE	PORTO	SANTAREM	SETUBAL	V.CASTELO	V.REAL	VISEU
a_{55}	.130 (.0164)	.167 (.0183)	.147 (.0112)	.103 (.0337)	.139 (.0196)	.194 (.0206)	.121 (.0141)	.138 (.0154)	.127 (.0096)
a_{56}	.761 (.2490)	.476 (.3802)	.650 (.1363)	.637 (.1592)	.512 (.3260)	.331 (.3980)	-.095 (.2662)	.552 (.2323)	.876 (.1330)
a_{66}	.129 (.3108)	.506 (.4398)	.286 (.1586)	.501 (.6646)	.430 (.3893)	.687 (.4700)	1.199 (.3538)	.336 (.3033)	-.102 (.1792)
a_{67}	.504 (.1592)	.471 (.2416)	.730 (.1316)	.847 (.4411)	.763 (.2235)	.650 (.3803)	.173 (.1847)	.529 (.1663)	.507 (.1040)
a_{77}	.322 (.2357)	.547 (.2550)	-.067 (.2008)	.198 (.5376)	.102 (.2709)	.363 (.4030)	.780 (.3464)	.391 (.2154)	.148 (.1895)
a_{78}	.462 (.1347)	.545 (.1682)	.589 (.1505)	.718 (.2819)	.464 (.1455)	.743 (.1423)	.825 (.1827)	.410 (.1721)	.643 (.1306)
a_{88}	.464 (.1684)	.385 (.2044)	.334 (.1793)	.322 (.3577)	.447 (.1826)	.119 (.1081)	-.054 (.2431)	.513 (.2209)	.274 (.1552)
a_{89}	.974 (.1630)	.733 (.1621)	.588 (.1354)	.940 (.2583)	.727 (.1627)	.915 (.1479)	.804 (.2209)	.776 (.2190)	1.328 (.1203)
a_{99}	-.107 (.1950)	.201 (.1916)	.407 (.1503)	-.023 (.3002)	.208 (.1855)	.048 (.1659)	.174 (.2509)	.187 (.2418)	-.342 (.1249)
$a_{9,10}$	.263 (.1801)	.253 (.2305)	.312 (.1266)	.554 (.2575)	.211 (.2071)	.231 (.1812)	.414 (.1579)	.583 (.1600)	.223 (.0992)
$a_{10,10}$	.626 (.3172)	.709 (.3506)	.604 (.2090)	.677 (.4273)	.712 (.3629)	.670 (.3477)	.242 (.3577)	-.0631 (.3077)	.666 (.1815)
$a_{10,11}$	.758 (.2777)	.777 (.2364)	.739 (.1626)	.383 (.4590)	.896 (.3611)	.797 (.2718)	.570 (.3156)	.622 (.3381)	.942 (.2366)
$a_{11,11}$	.312 (.3139)	.315 (.2450)	.378 (.1756)	.522 (.503)	.060 (.4151)	.262 (.2971)	.595 (.3312)	.456 (.3987)	.126 (.2427)
R^2	.9985	.9990	.9994	.9956	.9983	.9987	.9981	.9987	.9992
$\bar{R}^2$	.9977	.9985	.9990	.9937	.9973	.9980	.9970	.9981	.9989
F	1237.18	1959.77	2704.94	487.21	969.08	1368.59	902.25	1490.92	2011.13
df	56	58	55	59	54	56	55	57	60

TABLE 9.11: Estimated Values for the δ_r Coefficients for the Restricted OLS Estimation Procedure (Method I)

COEFF	AVEIRO	BEJA	BRAGA	BRAGANCA	C.BRANCO	COIMBRA	EVORA	FARO	GUARDA
EDUC 1	.0067 (.0187)	-	-	-	.0345 (.0230)	-.0054 (.0201)	.0384 (.0263)	.0084 (.0433)	-.0139 (.0268)
PEDUC 1	-	.0273 (.0758)	.0148 (.0354)	-	.0414 (.0798)	.0687 (.0493)	-	.0564 (.0549)	.0690 (.0445)
PCAP 1	.0279 (.0175)	-.0977 (.1306)	.0276 (.0304)	-.0083 (.0332)	.0091 (.0304)	.0086 (.0362)	.0083 (.0447)	-.0405 (.0666)	.0163 (.0256)
COST 1	-	.0143 (.0822)	.0910 (.0481)	.0246 (.0504)	.0361 (.0798)	.0130 (.0603)	-.0392 (.0840)	-	-
TEARN 1	-	-	-.0047 (.0157)	-	-	-.0069 (.0209)	-.0406 (.0406)	-	-
PUTEA 1	-	.0709 (.1370)	-	-	.0412 (.1003)	-	-.0194 (.0573)	.0469 (.1141)	-.0054 (.0265)
PUCLASS 1	-.0219 (.0350)	-.0389 (.0854)	.0518 (.0523)	-	-.0171 (.0312)	-	-.0678 (.0720)	-	-
BUS 1	-.0424 (.0562)	-	-	-	-	-	-.0116 (.0574)	-	-
HELP 1	-.0500 (.0397)	.0217 (.0837)	-	-.0467 (.0638)	-	-.0158 (.0367)	-	-	-.0816 (.0517)
UNQUAL 1	-	-	.0348 (.0579)	-	-.0477 (.1256)	.0321 (.0799)	-	.0087 (.0793)	-
GDP 1	-	-	-	.0355 (.0832)	-	-	-	.0537 (.0593)	.0710 (.0294)
LIFE 1	-	-	-	.0785 (.0140)	-	-	.0469 (.0660)	-	-
ILLIT 1	-	-	-	-	-	-	-	-	-
UNEMP 1	.0433 (.0380)	-.0199 (.1401)	.0144 (.0300)	-	-.0164 (.1078)	-	-	-.0238 (.0895)	-.0200 (.0299)
EARN 1	.0203 (.0300)	-.1209 (.1180)	-.0876 (.0573)	-.0274 (.0365)	-.0562 (.0354)	-.0378 (.0640)	.0226 (.0774)	-.0574 (.0704)	-
LFLEV 1	.1017 (.0740)	.1561 (.1389)	-	-	-	-	-	-	-
POPLEV 1	-	-.0524 (.1221)	-	-.0566 (.1036)	-.0106 (.0642)	.0467 (.0879)	-	-	-

TABLE 9.11: Estimated Values for the δ_r Coefficients for the Restricted OLS Estimation Procedure (Method T) (Continued)

COEFF	LEIRIA	LISBOA	PORTALEGRE	PORTO	SANTAREM	SETUBAL	V.CASTELO	V.REAL	VISEU
EDUC 1	.0213 (.0350)	.0162 (.0324)	-	-.0374 (.0264)	.0969 (.1670)	.0513 (.0400)	-.0033 (.0178)	.1418 (.0203)	-
PEDUC 1	-	-	-.0066 (.0647)	-	-.1651 (.1657)	.0236 (.0711)	-	-.0493 (.0508)	-
PCAP 1	.0042 (.1300)	-	-.0199 (.0321)	-	-.0115 (.0762)	.0220 (.0478)	.0431 (.0629)	-	.0116 (.0141)
COST 1	-	-	-	-.0631 (.0787)	-	-	-	-.0212 (.0482)	.0300 (.0092)
TEARN 1	-.0179 (.0228)	.0134 (.0248)	-	.0646 (.0515)	.0101 (.0535)	-.0832 (.0345)	-	-.0020 (.0114)	-
PUTEA 1	.0189 (.0588)	.0237 (.0876)	-	-.0917 (.0995)	.1900 (.4314)	-	-	.0187 (.0386)	-
PUCLASS 1	-	-	-.0186 (.0242)	-.1239 (.0731)	.0274 (.0987)	.0867 (.1791)	.0357 (.0709)	-.0189 (.0247)	-.0412 (.0191)
BUS 1	-	-	.0763 (.0877)	-	.1479 (.2435)	-	.0222 (.0450)	-	.0356 (.0606)
HELP 1	-.0398 (.0581)	-.0123 (.0517)	-.0213 (.0485)	-	-	-.1024 (.1167)	-.1307 (.1245)	-	-.0288 (.0402)
UNQUAL 1	-	.0188 (.0707)	.0492 (.1860)	-	-	-	-	-	-.0185 (.0482)
GDP 1	-.0672 (.0933)	-	.1373 (.1542)	-	-	-	.0403 (.0860)	-	.0026 (.0235)
LIFE 1	-	-	-	-	-	-.0274 (.1713)	-	-	-
ILLIT 1	-	-	-	-.2258 (.1677)	-	-	-	-	-
UNEMP 1	-	.0190 (.0577)	-.0355 (.0459)	-	.1417 (.1802)	-	.0769 (.0885)	.0558 (.0479)	-
EARN 1	-	-.0245 (.0272)	-	-	.0655 (.1375)	-	-.0066 (.0506)	-.0190 (.0338)	-
LFLEV 1	.0568 (.1309)	-	-	-	.0258 (.0980)	-	.0376 (.1092)	.0646 (.0475)	-
POPLEV 1	.0944 (.1250)	.0430 (.0638)	-.0683 (.1071)	-.2021 (.1696)	-	.0525 (.0679)	-	-	-

TABLE 9.12: Estimated Values for the $\underline{\delta}_p$ Coefficients for the Restricted OLS Estimation Procedure (Method I)

COEFF	AVEIRO	BEJA	BRAGA	BRAGANCA	C.BRANCO	COIMBRA	EVORA	FARO	GUARDA
EDUC 2	-.0064 (.0211)	-	.0170 (.0152)	-	-.0525 (.0230)	.0203 (.0518)	-.0044 (.0286)	-	-
PEDUC 2	-	-	-	-	-	-.1539 (.0917)	-	-.1642 (.1497)	-
PCAP 2	-	-	-	-	-.0516 (.0349)	-.3208 (.0745)	-.0776 (.0559)	-.1120 (.0811)	-.0693 (.0410)
COST 2	-	-.1086 (.1077)	-.0351 (.0453)	-.0232 (.0491)	-	-.0090 (.0435)	.1164 (.0641)	-	-
TEARN 2	-	.0260 (.0236)	-.0435 (.0174)	.0643 (.0214)	.0259 (.0219)	-.0309 (.0513)	-	.0086 (.0191)	.0607 (.0097)
PUTEA 2	.0450 (.0345)	-.0191 (.0475)	-.0274 (.0516)	.0092 (.0353)	-	.0243 (.0089)	-	-	.0433 (.0330)
PUCLASS 2	.0224 (.0385)	.0258 (.1073)	-.0484 (.0423)	-.0276 (.0322)	.0118 (.0242)	-.0401 (.1062)	-	-.0241 (.0352)	.0277 (.0143)
BUS 2	-	.0836 (.0841)	-	-.2136 (.0917)	-.1054 (.0866)	-	-.2042 (.0836)	-	-
HELP 2	-	-	-	.1452 (.0707	-	.1152 (.1349)	.3129 (.0668)	.3277 (.0906)	-
UNQUAL 2	-.0750 (.0486)	.0825 (.0669)	-.0977 (.0943)	-	-	-	-	.0162 (.0991)	.2057 (.0917)
GDP 2	-.0728 (.0448)	.0083 (.0775)	.1887 (.0610)	-.0454 (.0861)	-	-	-	.0597 (.0245)	-.0222 (.0691)
LIFE 2	.6612 (.2553)	-	-	.3023 (.2160)	-	-	-	-.2811 (.3893)	-
ILLIT 2	-	-	-	-	-.3657 (.2925)	-	-	-	-.7098 (.3206)
UNEMP 2	-.3118 (.0630)	-.0556 (.1054)	-.1129 (.0371)	-	-.1087 (.0633)	-	.0666 (.0550)	-.049 (.1581)	-
EARN 2	-	.1077 (.0790)	.1747 (.0538)	-	-	.0872 (.1010)	-.2239 (.0879)	.0158 (.1410)	-.1460 (.0520)
LFLEV 2	-.4210 (.1447)	-	-.2737 (.1037)	-	.0350 (.1507)	-	-	-	.1858 (.1355)
POPLEV 2	-	-	-	-.2078 (.1559)	-.2847 (.1430)	-	-.1607 (.0752)	-	-

TABLE 9.12: Estimated Values for the δ_r Coefficients for the Restricted OLS Estimation Procedure (Method I) (Continued)

COEFF	LEIRIA	LISBOA	PORTALEGRE	PORTO	SANTAREM	SETUBAL	V.CASTELO	V.REAL	VISEU
EDUC 2	-.0326 (.0383)	.0113 (.0366)	-.0189 (.0761)	-	-.1926 (.1251)	-	-.0419 (.0225)	.0270 (.0199)	.0541 (.0147)
PEDUC 2	.1268 (.0956)	-.0283 (.0174)	-	-.0689 (.0657)	.3048 (.1817)	.2785 (.2959)	-	-	-
PCAP 2	-	-.0581 (.0471)	.0151 (.0340)	-	-	-	-	-	-
COST 2	-.0976 (.1104)	.0084 (.0197)	.0334 (.0845)	.0573 (.0956)	-.1234 (.1994)	-	-.0448 (.0578)	-.0129 (.0510)	-
TEARN 2	-.0124 (.0255)	-.0639 (.0349)	-.0739 (.0907)	-.0308 (.0486)	.0338 (.0453)	-.1309 (.0492)	.0036 (.0207)	-	-
PUTEA 2	.0182 (.1199)	-	-	-	-.3112 (.2354)	.0579 (.1772)	-.1985 (.1303)	.1182 (.0411)	.0701 (.0303)
PUCLASS 2	.0516 (.0238)	-	.0475 (.0280)	.0503 (.0652)	-.0243 (.0762)	-	-.1064 (.0731)	.0309 (.0252)	.0865 (.0205)
BUS 1	-	-	-.0956 (.1148)	-	-	.3173 (.2326)	-	-	-.0822 (.0711)
HELP 2	-	.1042 (.0303)	-	.0447 (.0910)	-.0593 (.1701)	.0375 (.0697)	-	.0394 (.0396)	.0486 (.0387)
UNQUAL	-.2001 (.0770)	.0568 (.0639)	-	-	-	-	-.2270 (.4842)	-	-.0690 (.0491)
GDP 2	-	-	-.0524 (.3500)	-	-	-	.0830 (.1110)	.1989 (.1126)	.0117 (.0315)
LIFE 2	-	-	-	-	-	.7483 (.6977)	.1415 (.4017)	-	-
ILLIT 2	-	-	-	-	-	2.7211 (2.1118)	-	-	-
UNEMP 2	-.1813 (.0956)	-	-.2024 (.3251)	-.0376 (.0678)	-.3970 (.1797)	.2179 (.1462)	-.1723 (.0943)	-	-
EARN 2	-	-	.2419 (.1862)	-	-.0744 (.0909)	.2041 (.1342)	-	.1265 (.0416)	-
LFLEV 2	.0837 (.2201)	-	-	-.1160 (.1814)	-.0399 (.1753)	-	-	-.1711 (.1280)	-
POPLEV 2	-.1962 (.1681)	-	.1278 (.3254)	.0591 (.1621)	-	.1404 (.1226)	-.4698 (.1690)	-	-

estimates for some of the transition probabilities, with high standard errors and consequently large confidence intervals and non-significant t-values at the 5% level, the time-patterns of these transition probability estimates are the ones that fit better the patterns of the observed point estimates for the whole country. Therefore, for a better understanding of the influence of the explanatory variables in the changes of the transition probabilities by district, Tables 9.10 to 9.13 display the estimates of the different coefficients after performing the restricted OLS stepwise estimation procedure for each district individually. It is apparent from Tables 9.11 and 9.12 that the coefficients of the explanatory variables do not present the same sign for all districts, suggesting therefore, that any identification of the meanings of these coefficients should be made very carefully. If one district reveals that the change in one explanatory variable has a certain effect in the change of the transition probabilities, another district can present a completely opposite effect when changing the same explanatory variables. Therefore, when trying to draw conclusions about the causes of district disparities, one is faced with some less convincing results. The meanings of the behavioural relationships are, in some cases, not very clear. Measures such as reducing or making up the effects of distance (school transport, boarding), allowing the students to carry on studies by providing scholarships, equalising the qualification level of teachers, equalising the physical resources in order to create identical geographical conditions to the students, seem, however, to be attempts to ameliorate the regional differences.

9.3. Tentative Conclusions

At a national level, and with respect to both promotion and repetition probabilities, the extended Markov model reveals that an increase in the level of the average earnings (EARN), and an increase in the benefits given to the students by the social services (HELP) lead to a decrease in the national repetition probabilities, together with an increase in the national promotion probabilities. Furthermore, the promotion probabilities are also responsive to changes in the number of the unemployed workers (UNEMP) as well as to changes in the pupil-teacher ratio (PUTEA). This last explanatory variable, however, while it is not significant for Region 1, is seen to be quite important at national level and for Region 2.

It has been noted that the external factors are either on the demand side, and therefore difficult to control, or factors of supply, which are more open to intervention. Nevertheless, obscure interactions between supply and demand exist and must also be considered. This suggests that the analysis introduced in this study could be usefully further explored, either by trying to understand the effect of the supply side or the demand side factors separately upon the transition probabilities, or by including in the model a small set of explanatory variables at a time. Also, behavioural relationships describing the effect of these explanatory variables on the changes in the drop-out probabilities should be further analysed.

Chapter 10

CONCLUSIONS

The mathematical models developed during the 1960s and early 1970s to assist educational planners and decision-makers in the preparation of their educational plans have proved to be inefficient. The constant coefficient Markov chain is one such model, its performance at predicting school enrolment throwing into relief the limitations of the traditional Markov chain approach.

Although the Markov model was widely used by the governments as a basis for planning of operations to achieve educational policy objectives, only a few studies have thoroughly tested the model over a period of time to determine whether it is really valid for prediction purposes. Furthermore, all models described in the literature assume constant transition probabilities over time. However, if the model cannot describe with a reasonable confidence the historical trend, the values predicted for the future cannot reasonably be accepted.

The present study has started, therefore, by testing the stationary Markov chain model using a twelve year period data for a subsystem (basic preparatory and secondary levels) of the Portuguese educational system, the model being applied to the whole country, and to each district into which the country is administratively divided. The basic Markov model presented in Chapter 3 describes the theoretical framework of a stationary Markov model and the methods of estimating the transition probabilities. All the reasonable ways by which the transition probabilities could be estimated, under the assumption of being truly constant over time, have been tested in Chapters 4-5. The data used for this study show, however, that during the period of analysis there were strong fluctuations in the observed point estimates of the transition probabilities. Assuming that the disturbances observed in the data are due to the return of students from the old colonies Angola and Mozambique after the revolution that took place in April 1974, an iterative process was applied to separate these students from the observed data and to obtain an adjusted data matrix. Table 10.1 describes the least squares estimation procedures performed in both situations.

Table 10.1 Methods Used to Estimate the Constant Transition Probabilities for the Basic Markov Model

Original Data	Smoothed Data
Unrestricted OLS	Unrestricted OLS
Restricted OLS	Restricted OLS
Unrestricted GLS	

As expected, the traditional model has proved to be inappropriate as the transition probability estimates obtained are biased and non-efficient, with non-significant t-values at the 5% level and correspondingly large 95% confidence intervals having occurred for most of the estimates. The comparison between the different estimates obtained for the transition probabilities (see

Table 4.19) shows that there is little difference between the results presented, so that one cannot infer that one method is preferred in the sense that it gives a better proxy for the transition probabilities. The model has proved, therefore, to be insufficiently flexible to represent adequately the behaviour of school enrolment, the non-stationarity of the transition probabilities being apparent not only for the whole country but also in all districts.

Assuming that the non-stationarity of the transition probabilities is due to causal factors, it seemed worthwhile to quantify such relationships. Thus, if the explanatory variables of the relationship can be projected, or are determined by the authorities, forecasts for the transition probabilities can be obtained, using the behavioural relationships instead of trends.

The primary purpose of the study has been, therefore, to extend the traditional Markov model by allowing flexibility in the parameters. Linear behavioural relationships have been included in the model, the transition probabilities being assumed to be functions of a set of socio-economic and institutional explanatory variables. The theoretical framework of the extended Markov model was developed in Chapter 6 and its adaptation to the case study described in Chapter 7. Following the normal practice in classifying the causal factors behind changes in the transition probabilities, the explanatory variables selected have been divided into supply-side factors and demand-side factors. The extended Markov model was then applied to the same subsystem of the Portuguese educational system and different estimation procedures have been performed to produce time-varying estimates of the transition probabilities (see Chapters 7-8).

It must be recalled, however, that there were data limitations in the study. The short time-series period of observations used in the study has made it impossible to establish separate behavioural relationships for each transition probability. Thus, in the application of the extended Markov model to the Portuguese educational system, it has been assumed that the behavioural relationships for the set of the repetition probabilities only differ by the constant term, that is, by the estimate of the corresponding repetition probability if no exogenous influence take place. An analogous assumption has been made for the behavioural relationships for the set of promotion probabilities. The fact that the same set of explanatory variables has identical effects in the changes of all repetition probabilities, and another set for all promotion probabilities, may be an oversimplification in the model. While institutional explanatory variables such as teacher-pupil ratio, pupil-classroom ratio, and school bussing, for example, could be argued to have a greater effect on students enrolled at lower educational levels, economic explanatory variables such as the unemployment rate or the average earnings might well be more significant for students enrolled in upper educational levels. One suggests, therefore, that in further studies behavioural relationships for each level of education should be estimated.

Also, the unavailability of data concerning the number of new entrants to the school system at the entrance of each level of education, with exception of the first grade, has made it impossible to include these values in the model. Therefore, large bias ensued for the transition probability estimates relating to terminal or first grades of levels of education, for most of the attempts performed either using the basic Markov model or using the extended

Markov model. Nevertheless, when the causal structure is embodied in the model, the patterns of these transition probability estimates describe reasonably well the patterns of the corresponding observed point estimates, implying, therefore, the reliability of the behavioural relationships estimated (with the exclusion of the constant term).

Because of the existence of a certain degree of multicollinearity between the explanatory variables, principal components analysis was performed on supply side and demand side variables separately, and new sets of exogenous variables were used to estimate the time-varying transition probabilities.

A summary of the different attempts to produce time-varying estimates of the transition probabilities is presented in Table 10.2. Contrary to expectations, the results obtained using principal components are less satisfactory than the results obtained when the OLS stepwise regression (unrestricted) was applied simultaneously to all stacked district data (Method III). This method was the one that has produced the most acceptable time-patterns for the estimates. However, in general the fit achieved using time-varying transition probabilities to describe the fluctuations of the corresponding point estimates is far better than in the assumption of stationary parameters in the traditional Markov chain model. The overall improvement in the "predictive power" due to the use of explanatory variables, though not very high, is still noticeable.

The extended Markov model presented in this study seems to be an useful and significant extension of the traditional Markov chain model. By relaxing the assumption of stationary parameters it has

Table 10.2 Summary of the Methods Used to Estimate the Time-Varying Transition Probabilities for the Extended Model

Method	Type	Description	Output
I	unrest. and rest.	OLS stepwise regression applied to the whole country and for each district individually, using all explanatory variables.	Distinct behavioural relationships for the whole country and for each district.
II	unrest. and rest.	OLS estimation procedure using the principal components as explanatory variables and applied to the whole country and to each district individually.	Distinct behavioural relationships for the whole country and for each district.
III	unrest. and rest.	OLS stepwise regression applied to all stacked districts using all explanatory variables.	Identical behavioural relationships for the whole country and districts.
IV	unrest. and rest.	OLS stepwise regression performed by regions (industrial and rural) separately, stacking the districts of each region and using all explanatory variables.	Identical behavioural relationships for the districts of a region; different behavioural relationships by region.
V	unrest. and rest.	OLS stepwise regression applied to all stacked districts of a region and performed at first on the supply side and on the demand side explanatory variables separately. Using the significant variables regression was performed for each region.	Identical behavioural relationships for the districts of a region; different behavioural relationships by region.

Table 10.2 Summary of the Methods Used to Estimate the Time-Varying Transition Probabilities for the Extended Model
(Continued)

Method	Type	Description	Output
VI	unrest.	Pooled cross-section time--series estimation procedure applied by region and using the significant explanatory variables selected in Method IV (unrestricted).	Similar to Method IV
VII	unrest.	Pooled cross-section time-series estimation procedure applied by region and using the significant explanatory variables selected in Method V (unrestricted).	Similar to Method V
VIII	unrest. and rest.	OLS estimation procedure stacking over district, using the different principal components as explanatory variables and including dummy variables.	Distinct behavioural relationships per district.

been possible to model the changes of the transition probabilities, producing thus more meaningful estimates of these probabilities.

The results of the study reveal that the pupil-teacher ratio and the facilities offered to the students by the social services are the supply side explanatory variables that exert a significant effect on the the promotion probabilities; the unemployment rate and the average earnings level were the demand side variables that revealed the strongest influence on the students' decision on pursuing their studies. These results are in accordance with studies presented in the literature, which have shown family income and the educational attainment of the family to be variables that both exert large positive effect on school activity.

The application of the extended Markov model at the regional level (industrialized/rural) has revealed that the repetition probabilities for the region with rural characteristics are very sensitive to changes in the supply side factors. All methods performed have shown, however, that the promotion probabilities are more sensitive to external factors than the repetition probabilities. This suggests that in further studies the repetition probability could be replaced by the drop-out probability for estimating the behavioural relationships.

The existing multicollinearity in the model is obviously unsatisfactory and certainly led to an inefficient estimation of the model parameters. The elimination of this multicollinearity must be the subject of future work.

The analysis introduced in this study seems promising enough warrant to further developement, either by applying the extended Markov model to separate educational levels, or by trying to understand the effect of the supply side or the demand side factors separately upon the transition probabilities. For a better understanding of the effect of the explanatory variables on the transition probabilities, the behavioural relationships for the drop-out probabilities should also be analysed. Finally, regional analyses of the different behavioural relationships require further study, as they can be of great importance as support to educational planners and decision-makers in the preparation of the educational measures to be applied to each region or to each district.

BIBLIOGRAPHY

ABATENA, H. and TEDDLIE, C. (1986). "Rural school planning in Louisiana: a comparative study and an application of theoretical approaches". Socio-Economic Planning Sciences, Vol.20, 261-269.

ADELMAN, I. (1966). "A linear programming model of educational planning: a case study in Argentina". In ADELMAN, I. and THORBEKE, E. (Eds.), The Theory and Design of Economic Development. The Johns Hopkins Press, Baltimore, 385-417.

AITKEN, A.C. (1934). "On least-squares and linear combinations of observations". Proceedings of the Royal Society, Vol.55, 42-48.

AMEMIYA, T. (1981). "Qualitative response models: a survey". Journal of Economic Literature, Vol.19, 1483-1536.

AMES, E. and REITER, S. (1961). "Distributions of correlation coefficients in economic time-series". Journal of the American Statistical Association, Vol.56, 637-656.

ARMITAGE, J. and SABOT, R. (1987). "Socio-economic background and the returns of schooling in two low-income economies". Economica, Vol.54, 103-108.

ARMITAGE, P.H. and SMITH, C. (1967). "The development of computable models of the British educational system and their possible uses". In OECD (Ed.), Mathematical Models in Educational Planning, O.E.C.D., Paris, 159-197.

ARMITAGE, P.H., PHILLIPS, C.M. and DAVIES, J. (1970). "Towards a model of the upper secondary school system". Journal of the Royal Statistical Society (A), Vol.133, 166-205.

ARROW, K. (1973). "Higher education as a filter". Journal of Public Economy, Vol.2, 193-216.

AVERY, R.B. (1977). "Error components and seemingly unrelated regressions". Econometrica, Vol.45, 199-209.

BALTAGI, B.H. (1980). "On seemingly unrelated regressions with error components". Econometrica, Vol.48, 1547-1550.

BARTHOLOMEW, D.J. (1967,1973,1982). Stochastic Models for Social Processes (Editions 1,2,3), John Wiley & Sons, New York.

BARTHOLOMEW, D.J. (1981). Mathematical Methods in Social Sciences. Walter Lederman, London.

BARTHOLOMEW, D.J. and FORBES, A.F. (1979). Statistical Techniques for Manpower Planning. John Wiley & Sons, New York.

BENNET, R.J. (1980). Statistical Forecasting Concepts and Techniques in Modern Geography, Geoabstracts, 28. University of East Anglia Norwich.

BLAUG, M. (1967). "Approaches to educational planning". Economic Journal, Vol.77, 262-287.

BLAUG, M. (1972). An Introduction to Economics of Education. Penguin Books, Middlesex.

BLAUG, M. (1976). "The empirical status of human capital theory: a slightly jaundiced survey". Journal of Economic Literature, Vol.14, 827-855.

BLAUG, M. (1981). Education and the Employment Problem in Developing Countries. International Labour Office, Geneva.

BLAUG, M. (1983). Where Are We Now in the Economics of Education? A Special Professorial Lecture, University of London Institute of Education.

BLAUG, M., DOUGHERTY, C. and PSACHAROPOULOS, G. (1982). "The distribution of schooling and the distribution of earnings: ROSLA in 1972". Manchester School of Economic and Social Studies, Vol.50, 24-40.

BOARDMAN, A. (1975). Simultaneous Equations Models of the Educational Process. PhD Thesis, Carnegie-Mellon University.

BOARDMAN, A., DAVIS, A., SANDAY, R. (1977). "A simultaneous equations model of the educational process". Journal of Public Economics, Vol.7, 23-49.

BOWLES, S. (1967a). "A linear programming model of the educational sector". In BLAUG, M. (Ed.), Economics of Education, Vol.2, Penguin Books, Middlesex, 168-201.

BOWLES, S. (1967b). "The efficient allocation of resources in education". Quarterly Journal of Economics, Vol.81, 3, 189-219.

BOWLES, S. (1969). Planning Educational Systems for Economic Growth. Harvard University Press, Cambridge Massachusetts.

BOWLES, S. (1972). "Schooling and inequality from generation to generation". Journal of Political Economy, Vol.80, 219-251.

BRAY, J. (1965). "Economic modelling with computers". New Scientist, 22 July.

BRIDGES, R., JUDD, C. and MOOCK, P. (1979). The Determinants of Educational Outcomes. Ballinger, Cambridge.

BRITNEY, R.R. (1975). "Forecasting educational enrolments: comparison of a Markov chain and circuitless flow network model". Socio-Economic Planning Sciences, Vol.9, 53-60.

BROWN, D.J. (1975). "A smoothing solution to the school district enrolment projection problem". Educational Planning, Vol.2, 13-26.

BROWN, L.A. (1970). "On the use of Markov chains in movement research". Economic Geography, Vol.46, 393-403.

BRUNO, J.E. and NELKIN, I. (1975). "Logit methodology in educational policy and planning". Educational Planning, Vol.2, 27-35.

CARNEIRA, A.J.V.M. (1975). Higher Education Planning Problems at the Regional Level. The Case Study of the Northern Planning Region, Portugal. MPhil Thesis, University of Edinburgh.

CARRILHO, M.J. (1980). Tabuas Abreviadas de Mortalidade 1945-1975. Instituto National de Estatistica (INE), Lisbon.

CERVANTES, A.M.P.M. (1982). An Econometric Model for Wages and Enrolement in Education in Mexico with Forecasting. PhD Thesis, University of Pittsburg.

CHAN, M.W.L. (1981). "A Markovian approach to the study of the Canadian cattle industry". The Review of Economics and Statistics, Vol.63, 107-115.

CHOW, G.C. (1983). Econometrics. McGraw Hill, New York.

CLOUGH, D.J. and McREYNOLDS, W.P. (1966). "State transition model of an educational system incorporating a constraint theory of supply and demand". Ontario Journal of Educational Research, Vol.9, 1-18.

COHN, E. and KIKER, B.F. (1986). "Socio-economic background schooling, experience and monetary rewards in the United States". Economica, Vol.53, 497-503.

COLANDER, D.C. and GUTHRIE, R.S. (1981). "Great expectations: what the dickens do 'rational expectations' mean?". Journal of Post Keynesian Economics, Vol.3, 219-234.

COLLINS, L. (1974). "Estimating Markov transition probabilities from micro-unit data". Applied Statistics, Vol.23, 355-370.

COLLINS, L. (1975). An Introduction to Markov Chain Analysis. Concepts and Techniques in Modern Geography, 1, University of East Anglia, Norwich.

CONLISK, J. (1969). "Determinants of school enrollment and school performance", The Journal of Human Resources, Vol.4, 140-157.

CONLISK, J. (1976). "Interactive Markov chains". Journal of Mathematical Sociology, Vol.4, 157-185.

CORREA, H. (1963). The Economics of Human Resources. North-Holland Publishing Co., Amsterdam.

CORREA, H. and TINBERGEN, J. (1962). "Quantitative adaptation of education to acelerated growth". KYKLOS, Vol.15, 776-786.

DANTZIG, G.B. and ORDEN, A. (1953). "Duality theorems". RAND Report R.M.-1265, The RAND Corporation, Santa Monica, California.

DENISON, E. (1962). The Sources of Economics Growth in the United States and the Alternatives Before Us, Committee for Economic

Development, New York.

DENT, W. (1972). "A note on transition probability estimation". Australian Journal of Statistics, Vol.14, 217-221.

DOEBLIN, W. (1937). "Sur les proprietes asymptotiques de mouvement regis par certains types de chaines simples". Bulletin Mathematique de la Societe Roumaine de Sciences, Vol.39, 57-115.

DOEBLIN, W. (1940). "Elements d'une theorie generale des chaines simples constantes de Markoff". Annales Scientifiques de l'Ecole Normale Superieure, Vol.57, 61-111.

DOOB, J.L. (1942). "Topics in the theory of Markoff chains". Transactions of American Mathematical Society, Vol.52, 37-64.

DOOB, J.L. (1945). "Markoff chains - denumerable case". Transactions of American Mathematical Society, Vol.58, 455-473.

DORN, W.S. (1960). "Duality in quadratic programming". Quarterly Journal of Applied Mathematics, Vol.18, 155-162.

DRAPER, N.R. and SMITH, H. (1981). Applied Regression Analysis. John Wiley and Sons, New York.

DUFOUR, J.M. (1984). "Unbiasedness of predictions from estimated autoregressions when the true order is unknown". Econometrica, Vol.52, 209-215.

DURBIN, J. (1984). "Present position and potential developments: some personal views - time-series analysis". Journal of the Royal Statistical Society (A), Vol.147, 161-173.

ECKAUS, R.S., SAFTY, A., AHMADEL,H. NORMAN, V.D. (1974). "An appraisal of the calculations of rates of return to higher education", in GORDON, M.S. (Ed.), Higher Education and the Labour Market. McGraw Hill, New York.

EIDE, K. (1983). "New features in educational planning". In OECD (Ed.), Educational Planning-a Reappraisal. OECD, Paris.

FAREBROTHER, R.W. (1980). "The Durbin-Watson test for serial correlation when there is no intercept in the regression". Econometrica, Vol.48, 1553-1563.

FARRAR, D.E. and GLAUBER, R.R. (1967). "Multicollinearity in regression analysis: the problem revisited". Review of Economics and Statistics, Vol.49, 92-107.

FENSKE, R.H. and SCOTT, C.S. (1973). "College students' goals, plans and background characteristics: a synthesis of three empirical studies". Research in Higher Education, Vol.1, 101-118.

FISHER, D. and KNESPER, D. (1983). "Markov models and the utilization of Mental Health Services: a study of endogenously depressed patients". Socio-Economic Planning Sciences, Vol.17, 21-31.

FOLK, M. (1976). "A critique of some future prediction techniques and their implications for educational planners". Educational

Planning, Vol.2, 35-52.

FREEDMAN, D. (1971). Markov Chains. Holden Day, San Francisco.

FREIRE, M.E.S. (1979). The Economic Value of Education in Portugal: Some Comments on the Relations Between Education and Economy. Fundacao Calouste Gulbenkian, Lisbon (mimeographed).

FREIRE, M.E.S. and SILVA, J.J.R.F. (1974). Study of the Portuguese Educational System Using a Markov Chain Model. Universidade Nova de Lisboa, Lisbon.

GANESH, L.S. (1986). "Identifying temporal relationships in educational planning". Socio-Economic Planning Sciences, Vol.20, 233-236.

GANI, J. (1963). "Formulae for projecting enrolments and degrees awarded in Universities". Journal of Royal Statistical Society (A), Vol.126, 400-409.

GINSBERG, R.B. (1972a). "Critique of probabilistic models: application of the semi-Markov model to migration". Journal of Mathematical Sociology, Vol.2, 63-82.

GINSBERG, R.B. (1972b). "Incoporating causal structure and exogenous information with probabilistic models: with special reference to choice gravity, migration and Markov chains". Journal of Mathematical Sociology, Vol.2, 83-103.

GIRARDIEAU, C. (1967). "The introduction of vocational training into a model of optimal allocation of resources between the economy and the educational system". Bulletin du CEPREL, No.8, 75-117.

GODET, M. (1977). Crise de la Prevision: Essor de la Prospective. Presses Universitaires de France, Paris.

GOKCE, D.R. (1981). Personal Distribution of Earnings and a Markov Chain Analysis of Chance Factors and Their Effects on Earnings Inequality. PhD Thesis, Georgia State University-College of Business Administration.

GOLDFELD, S.H. and QUANDT, R.E. (1972). Non-Linear Methods in Econometrics. North Holland Publishing, Amsterdam.

GOODMAN, L. (1961). "Statistical methods for the mover-stayer model". Journal of the American Statistical Association, Vol.56, 841-868.

GOODMAN, L. (1965). "On the statistical analysis of mobility Tables". American Journal of Sociology, Vol.70, 564-585.

GOODMAN, L. (1969). "How to ransack social mobility tables and other kinds of cross classification tables". American Journal of Sociology, Vol.75, 1-39.

GOODMAN, L. (1979). "Multiplicative models for the analysis of occupational probability tables and other kinds of cross-classification tables". American Journal of Sociology, Vol.84, 804-819.

GREMBOWSKI, D. (1984). "Causal models in plan evaluation". Socio-Economic Planning Sciences, Vol.18, 255-261.

HILAK, J. (1969). Facteurs Economiques et Sociaux dans la Determination de la Demande de Places par les Eleves. Les Principes de la Planification de l'Education: Conferences et discussions 33. IIPE/UNESCO, Paris.

HARNQVIST, K. (1978). Individual Demand for Education: Analytical Report. OECD, Paris.

HARRIS, C.S. (1980). "The credibility of educational forecasts". Journal of Teacher Education, Vol.31, 35-38.

HENDERSON, C.R. (1971). "Comment on the use of error components models in combining cross-section with time-series data". Econometrica, Vol.39, 397-401.

HENRY, N.W. (1971). "The retention model: a Markov chain with variable transition probabilities". Journal of the American Statistical Association, Vol.66, 264-267.

HOEL, P.G. (1954). "A test for Markoff chains". Biometrika, Vol.41, 430-433.

INE (1971,1972,...,1980). Estatistica da Educacao. Instituto Nacional de Estatistica (INE),Lisbon.

INE (1972). IX Census da Populacao. INE,Lisbon.

INE (1983). XI Census da Populacao. INE, Lisbon.

INE (1984). Estatistica da Educacao 1981-1983. INE, Lisbon.

JACOBI, E.G. (1959). Methods of School Enrolment Projection. Educational Studies and Documents, No.32, UNESCO, Paris.

JOHNSTON, J. (1972). Econometric Methods. Mc Graw Hill, New York.

JOHNSTONE, J.N. (1981). Indicators of Education Systems. Kogan Page, London/UNESCO, Paris.

JOHNSTONE, J.N. (1982). "Three useful but often forgotten education system indicators". Socio-Economic Planning Sciences. Vol.16, 163-166.

JOHNSTONE, J.N. and PHILP, H. (1973). "The application of a Markov chain in educational planning". Socio-Economic Planning Sciences, Vol.7, 283-294.

JOHNSTONE, J.N. and PHILP, H. (1975). "Letters to the editor on JOHNSTONE and PHILP's Markov educational planning model". Socio-Economic Planning Sciences, Vol.9, 89-91.

JUDD, C.H. and KENNY, D.A. (1981). Estimating the Effects of Social Interventions. Cambridge University Press, Cambridge.

KAISER, H.F. (1960). "The application of electronic computers to factor analysis". Educational and Psychological Measurement, Vol.20, 140-151.

KAKWANI, N.C. (1967). "The unbiasedness of Zellner's seemingly unrelated regression equations estimators". Journal of the American Statistical Association, Vol.62, 141-142.

KAYSER, B.D. (1976). "Educational aspirations as a Markovian process: a panel study of college and non-college plans during high-school". Journal of Mathematical Sociology, Vol.4, 295-306.

KEMENY, J.G. and SNELL, J.L. (1967). Finite Markov Chains. D. Van Nostrand Co., Princeton, New Jersey.

KMENTA, A.J. (1971). Elements of Econometrics. Macmillan, London.

KMENTA, J. and GILBERT, R.F. (1968). "Small sample properties of alternative estimators of seemingly unrelated regressions". Journal of the American Statistical Association, Vol.63, 1180-1200.

KOGAN, M. (1980). The Relationship Between Education Policy, Planning and Politics. OECD, Paris (mimeographed).

KOUTSOYIANNIS, A. (1973). Theory of Econometrics. Macmillan, London.

KUHN, H. and TUCKER, A. (1951). "Non-linear programming". In NEYMAN, J. (Ed.), Proceedings of the Second Berkeley Symposium. The University of California Press, Berkeley, 481-492.

KUMAR, R.C., PARDO, S. and WATSON, C. (1980). "School dropouts in Ontario: a simulation of the phenomenon in a local school system". Socio-Economic Planning Sciences, Vol.14, 217-231.

LEE, T.C., JUDGE, G.G., ZELLNER, A. (1970). Estimating the Parameter of the Markov Probability Model from Aggregate Time-Series Data. North-Holland Publishing Co., Amsterdam.

LEIBOWITZ, A. (1974). "Education and home production". American Economic Review, Vol.64, 243-250.

LERMAN, R. (1974). "Some determinants of youth school activity". The Journal of Human Resources, Vol.7, 366-383.

McCALL, J. (1971). "A Markovian model of income dynamics". Journal of the American Statistical Association, Vol.66, 439-447.

McFARLAND, D.D. (1970). "Intragenerational social mobility as a Markov process: including a time-stationary Markovian model that explains observed declines in mobility rates over time". American Sociological Review, Vol.35, 463-475.

McGINNIS, R. (1968). "A stochastic model of social mobility". American Sociological Review, Vol.33, 712-721.

McKELVEY, R.D. and ZAVOINA, W. (1975). "A statistical model for the analysis of ordinal level dependent variables". Journal of Mathematical Sociology", Vol.4, 103-120.

MaCRAE, E.C. (1977). "Estimation of time-varying Markov processes with aggregate data". Econometrica, Vol.45, 183-197.

MACE, J. (1984). "The economics of education: a revisionist's view". Higher Education Review, Summer, 39-56.

MADANSKY, A. (1959). "Least squares estimation in finite Markov processes". Psychometica, Vol.24, 137-144.

MADDALA, G.S. (1971). "The use of variance component models in pooling cross-section and time-series data". Econometrica, Vol.39, 341-382.

MADDALA, G.S. (1977). Econometrics. McGraw Hill, New York.

MARSHALL, F.R., BRIGGS, V.M. and KING, A.G. (1984). Labor Economics = Wages, Employment and Trade Unionism. Irwin Inc., Homewood, Illinois.

MASTERS, S.H. (1969). "The effect of family income on children's education: some findings on inequality of opportunity". The Journal of Human Resources, Vol.4, 158-175.

MATTILA, J.P. (1982). "Determinants of male enrolments: a time-series analysis". Review of Economics and Statistics, Vol.64, 242-251.

MELLOR, C. and HESSNER, C. (1986). "Structural changes in the yellow fats market". Applied Economics, Vol.18, 663-673.

MILLER, G.A. (1952). "Finite Markov processes in psychology". Psychometrika, Vol.17, 149-167.

MOORE, P. (1983). "Higher education: the next decade". Journal of the Royal Statistical Society (A), Vol.146, 213-245.

NERLOVE, M. (1971). "Further evidence on the estimation of dynamic economic relations from a time-series of cross-sections". Econometrica, Vol.39, 359-382.

NICHOLLS, D.F. and PAGAN, A.R. (1977). "Specification of the disturbance for efficient estimation - an extended analysis". Econometrica, vol.45, 211-217.

OECD (1967a). A Technical Evaluation of the First Stage of the Mediterranean Regional Project. OECD, Paris.

OECD (1967b). Methods and Statistical Needs for Educational Planning. OECD, Paris.

OECD (1969). Efficiency in Resource Utilization in Education Technical Reports. OECD, Paris.

OECD (1970). S.O.M. a Simulation Option Model of the Educational System. OECD, Paris.

OECD (1973). Mathematical Models for the Education Sector: A Survey. OECD, Paris.

OECD (1974). Indicators of Performance of Educational Systems. OECD, Paris.

OECD (1975). Education, Inequality and Life Chances. OECD, Paris.

OECD (1979a). Individual Demand for Education: General Rapport and Case Studies: France, Germany, Greece and U. Kingdom. OECD, Paris.

OECD (1979b). Future Educational Policies in the Changing Social and Economic Context. OECD, Paris.

OECD (1979c). Education and Regional Development: General Report. OECD, Paris.

OECD (1983). Educational Planning - A Reappraisal. OECD, Paris.

OECD (1984). Reviews of National Policies for Education - Portugal. OECD, Paris.

ORR, L. (1972). "The dependence of transition proportions in the education system on observed social factors and social characteristics". Journal of the Royal Statistical Society (A), Vol.135, 74-95.

PANITCHPAKDI, D. (1974). Educational Growth in Developing Countries: an Empirical Analysis. Studies in Development and Planning, Vol.5. Rotterdam University Press, Rotterdam.

PARNES, H. (1960). Forecasting Educational Needs for Economic and Social Development. OECD, Paris.

PARNES, H. (1962). "Manpower analysis in educational planning". In BLAUG, M. (Ed.), Economics of Education, Penguin Books, Middlesex, Vol.2, 263-286.

PERRY, W.D. (1975). General Quantitative Models and Policy Analysis of Turnover and Attendance in Manpower Programs. PhD Thesis, School of Urban and Public Affairs, Carnegie - Mellon University.

PIDOT, G.B. (1969). "A principal component analysis of the determinants of local government fiscal patterns". Review of Economics and Statistics, Vol.51, 176-188.

PYNDICK, R.S. and RUBINFELD, D.L. (1981). Econometric Models and Economic Forecasts. McGraw Hill, New York.

PLOTNICKI, W.J. and GARFINKEL, R.S. (1986). "Scheduling academic courses to maximize student flow: a simulation approach". Socio-Economic Planning Sciences, Vol.20, 193-199.

PSACHAROPOULOS, G. (1973). Returns to Education: An International Comparison. Elsevier Scientific Publishing Co., London.

PSACHAROPOULOS, G. (1975a). Earnings and Education in OECD Countries. OECD, Paris.

PSACHAROPOULOS, G. (1975b). "The macro-planning of education: a clarification of issues and a look into the future". Comparative Education Review, Vol.19, 214-224.

PSACHAROPOULOS, G. (1981a). "Returns to education: an updated international comparison. Comparative Education. Vol.17, 321-341.

PSACHAROPOULOS, G. (1981b). "Education and the structure of earnings in Portugal". De Economist, Vol.129, 532-545.

PSACHAROPOULOS, G. (1982a). "An analysis of the determinants of the demand for upper secondary education in Portugal". Economics of Education Review, Vol.2, 233-251.

PSACHAROPOULOS, G. (1982b). "Educacao como investimento". Financas e Desenvolvimento, Sept., 39-42.

PSACHAROPOULOS, G. and SOUMELIS, C. (1979). "A quantitative analysis of the demand for higher education". Higher Education, Vol.8, 159-177.

RAO, B.L.S.P. (1972). "Maximum likelihood estimation for Markov processes". Annals of the Institute of Statistical Mathematics, Vol.24, 333-345.

ROBINSON, E.A.G. and VAIZEY, J.E. (1966). The Economics of Education. Macmillan, London.

SAOPEDRO, M.E., AMORIM, A., CALADO, A., GASPAR,I. (1983). Diagnostico/ Previsoes. GEP, Ministry of Education, Lisbon.

SCHULTZ, T.W. (1959). "Investment in man: an economist's view". Social Review, Vol.33, 109-117.

SCHULTZ, T.W. (1961). "Investment in human capital". American Economic Review, Vol.51, 1-17.

SCHULTZ, T.W. (1963). The Economic Value of Education. Columbia University Press, New York.

SEN, A.K. (1964). "Comments on the paper by Messrs Tinbergen and Bos" In OECD (Ed.). The Residual Factor and Economic Growth, OECD, Paris, 188-197.

SHIPMAN, W.A. (1978). An Econometric Study of Teenage Labour Market and School Enrolment Behaviour. PhD Thesis, University of Pittsburgh.

SHORROKS, A.F. (1976). "Income mobility and the Markov assumption". The Economic Journal, Vol.86, 566-578.

SOARES, M.C., SAOPEDRO, M.E. and MAGALHAES, M.M. (1980). Procura Social do Ensino: 6º ano, 9º ano and 11º ano, GEP/Ministry of Education, Lisbon.

SOUMELIS, C. (1981). "Do rates of return matter for public and private educational decisions?" Higher Education. Vol.10, 353-361.

SOUMELIS, C. (1983). "The evolution of educational planning concepts and approaches - an overview". In OECD (Ed.), Educational Planning - a Reappraisal, OECD, Paris, 12-51.

SPILLERMAN, S. (1972). "The analysis of mobility processes by the introduction of independent variables into a Markov chain. American Sociological Review, Vol.37, 277-294.

SPILLERMAN, S. (1973). "Extensions of the mover-stayer model". American Journal of Sociology, Vol.78, 599-626.

STONE, R. (1965). "Model of educational system". Minerva, Vol.3, 172-186.

STONE, R. (1972). "A Markovian education model and other examples linking social behaviour to the economy". The Journal of the Royal Statistical Society (A), Vol.135, 511-543.

SVENNILSON, I.,EDDING, F. and ELVIN, R. (1962). Targets for Education in 1970. OECD, Paris.

TAKAYAMA, A. (1974). Mathematical Economics. Dryden, Illinois.

TAKAYAMA, T. and JUDGE, G.G. (1971). Spatial and Temporal Price and Allocation Models. North-Holland Publishing Co., Amsterdam.

THEIL, H. (1971). Principles of Econometrics. John Wiley, New York.

THONSTAD, T. (1967). Mathematical Models in Educational Planning. OECD, Paris.

THONSTAD, T. (1969). Education and Manpower Theoretical Models and Empirical Applications. Oliver and Boyd, Edinburgh.

THONSTAD, T. (1981). Analysing and Projecting School Enrolments in Developing Countries: A Manual of Methodology. UNESCO, Paris.

THOOLEN, B.A. (1970). Some Problems of S.O.M. (e) Models. Dutch Central Planning Bureau, The Hague (mimeographed).

THORNDIKE, R.M. (1978). Correlation Procedures for Research. Gaidhere Press, New York.

TILAK, J.B.G. (1982). "Educational planning and the international economic order". Comparative Education, Vol.18, 107-121.

TINBERGEN, J. and BOS, H.C. (1965). "A planning model for the educational requirements of economic development". In BLAUG, M. (Ed.), Economics of Education, Vol.2, Penguin Book, Middlesex.

UNESCO (1955). World Survey of Education I: Handbook of Educational Organization and Statistics. UNESCO, Paris.

UNESCO (1966). An Asian Model of Educational Development: Perspectives for 1965-80. UNESCO, Paris.

UNESCO (1974a). The UNESCO Educational Simulation Model (ESM). UNESCO, Paris.

UNESCO (1974b). Social Indicators: Problems of Definition and of Selection. UNESCO, Paris.

UNESCO (1974c). A System of Educational Indicators. UNESCO, Paris.

UNESCO (1976a). Methods of Projecting School Enrolment in Developing Countries. UNESCO, Paris.

UNESCO (1976b). International Standard Classification of Education. UNESCO, Paris.

UNESCO (1981). Planning Education for Reducing Inequalities. UNESCO, Paris.

UNESCO (1983). Educational Planning in the Context of Current Development Problems. UNESCO/IIEP, Paris.

VAIZEY, J. (1962). The Economics of Education. Faber and Faber, London.

WALLACE, T.D. (1964). "Efficiences for step-wise regressions". Journal of the American Statistical Association, Vol.59, 1179-1182.

WALLACE, T.D. and HUSSAIN, A. (1969). "The use of error components models in combining cross section with time series data". Econometrica, Vol.37, 55-68.

WHITEMAN, R.E. (1985). Enrolment and Economic Impacts on Michigan Private Liberal Arts Colleges. PhD Thesis. Michigan State University.

WILLIAMS, G. (1979). "Educational planning, past, present and future". In Education Policy Bulletin, Winter. OECD, Paris. (mimeographed), 1-15.

WILLIAMS, G. (1983). "An agenda for the future: A framework for educational policy planning". In OECD (Ed.), Educational Planning - A Reappraisal, OECD, Paris, 334-360.

WOLFE, P. (1959). "The simplex method for quadratic programming". Econometrica, Vol.27, 382-398.

ZELLNER, A. (1962). "An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias". Journal of the American Statistical Association, Vol.57, 348-368.

ZELLNER, A. (1963). "Estimators of seemingly unrelated regressions: some exact finite sample results". Journal of the American Statistical Association, Vol.58, 977-992.

ZELLNER, A. and THEIL, H. (1962). "Three stage least squares: simultaneous estimation of simultaneous squations". Econometrica, Vol.30, 54-78.